

```

; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-617

```

```

Query Match      37.3%; Score 44; DB 3; Length 100;
Best Local Similarity 43.8%; Pred.No. 7.6;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      3 WNKTLQGLLDRIQAPP 18
Db      65 WAELEGLLQRLSRP 80

```

```

Search completed: November 10, 2004, 15:57:21
Job time : 11.1779 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 53.1601 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-68

Perfect score: 118

Sequence: 1 ATWMTLQGLLDRIQAFSSPH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	22	14	US-10-092-750-68
2	53	44.9	301	15	Sequence 68, Appl
3	51	43.2	403	9	Sequence 46367, A
4	50.5	42.8	112	10	Sequence 10282, A
5	50	42.4	113	14	Sequence 108, App
6	50	42.4	339	9	Sequence 3055, App
7	50	42.4	339	10	Sequence 246, App
8	50	42.4	486	15	Sequence 246, App
9	50	42.4	486	16	Sequence 2238, App
10	49.5	41.9	352	15	Sequence 1037, App
11	48	40.7	139	14	Sequence 42581, A
12	48	40.7	213	16	Sequence 2028, A
13	48	40.7	299	15	Sequence 115384, A
					Sequence 45536, A

ALIGNMENTS

RESULT 1

US-10-092-750-68
; Sequence 68, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-68

Query Match 100.0%; Score 118; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATWMTLQGLLDRIQAFSSPH 22

DB 1 ATWMTLQGLLDRIQAFSSPH 22

RESULT 2

US-10-282-122A-46367
; Sequence 46367, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

Sequence 174130,
Sequence 174132,
Sequence 171903,
Sequence 204400,
Sequence 138270,
Sequence 382, App
Sequence 565, App
Sequence 779, App
Sequence 168590,
Sequence 129534,
Sequence 57759, A
Sequence 346804,
Sequence 204192,
Sequence 255172,
Sequence 223390,
Sequence 275172,
Sequence 110986,
Sequence 984, App
Sequence 308284,
Sequence 43841, A
Sequence 287969,
Sequence 5059, App
Sequence 43511, A
Sequence 10325, A
Sequence 541, App
Sequence 203466,
Sequence 203468,
Sequence 3653, App
Sequence 170656,
Sequence 128047,
Sequence 197179,

14 48 40.7 496 16 US-10-437-963-174130
15 48 40.7 500 16 US-10-437-963-174132
16 48 40.7 974 16 US-10-437-963-171903
17 48 40.7 1565 16 US-10-437-963-204400
18 47.5 40.3 1513 16 US-10-437-963-138270
19 47 39.8 40 16 US-10-109-048-382
20 47 39.8 89 16 US-10-109-048-565
21 47 39.8 92 16 US-10-109-048-779
22 47 39.8 119 16 US-10-109-048-159
23 47 39.8 740 16 US-10-437-963-168590
24 47 39.8 982 16 US-10-437-963-129534
25 46.5 39.4 356 15 US-10-282-122A-57759
26 46 39.0 87 17 US-10-425-115-346804
27 46 39.0 110 16 US-10-437-963-204192
28 46 39.0 207 15 US-10-424-599-255172
29 46 39.0 440 15 US-10-424-599-223390
30 46 39.0 500 15 US-10-424-599-275172
31 46 39.0 1365 16 US-10-437-963-110986
32 45 38.1 158 9 US-09-925-300-984
33 45 38.1 246 17 US-10-425-115-308284
34 45 38.1 311 16 US-10-787-701-43841
35 45 38.1 341 17 US-10-425-115-287969
36 45 38.1 344 9 US-09-815-242-5059
37 45 38.1 425 17 US-10-282-122A-43511
38 45 38.1 403 14 US-10-739-930-10325
39 45 38.1 503 14 US-10-310-154-541
40 45 38.1 733 16 US-10-437-963-203466
41 45 38.1 868 16 US-10-437-963-203468
42 45 38.1 1035 14 US-10-389-493-3653
43 45 38.1 1417 16 US-10-437-963-170656
44 45 38.1 1452 16 US-10-437-963-128047
45 45 38.1 1555 16 US-10-437-963-197179


```
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; NUMBER OF SEQ ID NOS: 4996
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3055
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3055

Query Match          42.4%; Score 50; DB 14; Length 113;
Best Local Similarity 50.0%; Freq. No. 7.7;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 TWMKTLQGLLDRIQAF 17
      |||:||||:|:
Db      32 TWQQAULTGLLRNQTY 47

RESULT 6
US-09-981-876-246
; Sequence 246, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
```

```

; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/057,761
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,891
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: Patentin Ver. 2.0

```

```

; SEQ ID NO 246
; LENGTH: 339

Query Match 42.4%; Score 50; DB 9; Length 339;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 TWMTKLOGLDLRIQAF 17
   |||:||||:
Db 223 TWQALTGLLRMQTY 238

RESULT 7
US-09-148-545-246
; Sequence 246, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613

```

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 246
LENGTH: 339

Query Match 42.4%; Score 50; DB 10; Length 339;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TWMKTLQGLLDRIQAF 17
 Db 223 TWQALTGLLRMQTY 238

RESULT 8
 US-10-264-237-2238
 ; Sequence 2238, Application US/10264237
 ; Publication No. US2004009491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA131P1
 ; CURRENT APPLICATION NUMBER: US/10/264,237
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450
 ; PRIOR FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: US 60/205,515
 ; PRIOR FILING DATE: 2000-05-19
 ; NUMBER OF SEQ ID NOS: 2876
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 2238
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (276)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 US-10-284-237-2238

Query Match 42.4%; Score 50; DB 15; Length 486;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TWMKTLQGLLDRIQAF 17
 Db 365 TWQALTGLLRMQTY 380

RESULT 9
 US-10-408-765A-1037
 ; Sequence 1037, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1037
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-1037

Query Match 42.4%; Score 50; DB 16; Length 486;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TWMKTLQGLLDRIQAF 17
 Db 365 TWQALTGLLRMQTY 380

RESULT 10
 US-10-282-122A-42581
 ; Sequence 42581, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Chisen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 42581
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-10-282-122A-42581

Query Match 41.9%; Score 49.5; DB 15; Length 352;
 Best Local Similarity 40.0%; Pred. No. 32;
 Matches 10; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

Qy 2 TW-----MKTLQGLLDRIQAFPPSP 21
 Db 253 TWGLLGSMPTMEGTEDKLYAIPGSP 277

RESULT 11
 US-10-094-749-2028
 ; Sequence 2028, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO

```

; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2028
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2028

Query Match          40.7%; Score 48; DB 14; Length 139;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 7 LQGLLDRIQAFPPSPH 22
Db 6 MQGLLTPTHTAFPQPH 21

RESULT 12
US-10-437-963-115384
; Sequence 115384, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115384
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18987C.1.pdp
US-10-437-963-115384

Query Match          40.7%; Score 48; DB 16; Length 213;
Best Local Similarity 61.5%; Pred. No. 31;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LQGLLDRIQAFPS 19
Db 15 MQGLLDHVDAPFA 27

RESULT 13
US-10-282-122A-45536
```

```

; Sequence 45536, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45536
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (183)..(183)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (185)..(185)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-45536

Query Match          40.7%; Score 48; DB 15; Length 299;
Best Local Similarity 42.1%; Pred. No. 45;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MKTLQGLLDRIQAFPPSPH 22
Db 88 ISTTKGLIIOIQALPCKPH 106

RESULT 14
US-10-437-963-174130
; Sequence 174130, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

Fri Nov 12 14:55:56 2004

APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 174130
 LENGTH: 496
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_72100C.1.pep
 US-10-437-963-174130

Query Match 40.7%; Score 48; DB 16; Length 496;
 Best Local Similarity 66.7%; Pred. No. 78;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 8 QGLLDRIQAFPSPPH 22
 Db 176 QGLLDIIQAMPHDAH 190

RESULT 15
 US-10-437-963-174132
 Sequence 174132, Application US/10437963
 Publication No. US20040123343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 174132
 LENGTH: 500
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(500)
 OTHER INFORMATION: unsure at all Xaa locations
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_72102C.1.pep
 US-10-437-963-174132

Query Match 40.7%; Score 48; DB 16; Length 500;
 Best Local Similarity 66.7%; Pred. No. 79;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 8 QGLLDRIQAFPSPPH 22
 Db 157 QGLLDIIQAMPHDAH 171

Search completed: November 11, 2004, 07:41:45
 Job time : 53.1601 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 7.51601 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-68

Perfect score: 118
Sequence: 1 ATWMTLQGLLDRIQAPFSSPH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: PIR.79.*
- 2: PIR1.*
- 3: PIR2.*
- 4: PIR3.*
- 5: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	51	43.2	403	2 A91097	probable deacetyla
2	51	43.2	403	2 H65070	hypothetical prote
3	51	43.2	403	2 E85942	probable deacetyla
4	49	41.5	296	2 G87505	histone deacetylase
5	48	40.7	2225	1 A23443	pyrimidine synthet
6	47	39.8	603	2 S61504	glycogen(starch) s
7	46	39.0	503	2 B81690	probable sodium-tr
8	46	39.0	544	2 H86828	fructuronate reduc
9	46	39.0	732	1 A35655	peptidyl-dipeptida
10	46	39.0	1097	2 A2572	hypothetical prote
11	46	39.0	1298	2 T47523	DNA-binding protei
12	46	39.0	1312	1 A34171	peptidyl-dipeptida
13	45	38.1	344	2 B83605	probable acetylpol
14	45	38.1	480	2 T49157	citrate synthase-1
15	45	38.1	516	2 S32007	citrate synthase -
16	45	38.1	527	1 S25481	heat shock transcr
17	45	38.1	733	2 A83636	sulfite reductase
18	45	38.1	1313	1 JC2038	peptidyl-dipeptida
19	45	38.1	1616	2 I37183	gene APXL protein
20	45	38.1	3433	1 S28381	utrophin - human
21	44	37.3	246	2 A30589	hypothetical prote
22	44	37.3	269	2 J27536	chitinase (EC 3.2.
23	44	37.3	306	2 T2650	hypothetical prote
24	44	37.3	345	2 S64476	hypothetical prote
25	44	37.3	515	2 JC5458	inulinase (EC 3.2.
26	44	37.3	614	2 F96791	hypothetical prote
27	44	37.3	732	1 S05238	peptidyl-dipeptida
28	44	37.3	952	2 T48510	MVB like protein -
29	44	37.3	1306	1 A31759	peptidyl-dipeptida

ALIGNMENTS

RESULT 1

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

A91097

probable deacetylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A91097
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A95629, UID:21156231, PMID:11258796
A;Accession: A91097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <HAY>
A;Cross-references: UNIPROT:Q46805; GB:BA000007; PIDN:BA837168.1; PID:gl3363217; GSPDB:
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC53745
C;Superfamily: succinyl-diaminopimelate desuccinylase

30 44 37.3 4006 2 T09070
31 43.5 36.9 397 1 PC1219
32 43.5 36.9 514 2 T49158
33 43 36.4 175 2 S52931
34 43 36.4 203 2 G95309
35 43 36.4 235 2 AE1144
36 43 36.4 235 2 AE1503
37 43 36.4 356 2 G75444
38 43 36.4 424 2 E65088
39 43 36.4 424 2 H70501
40 43 36.4 456 2 D84202
41 43 36.4 723 2 AG0887
42 43 36.4 739 2 D91116
43 43 36.4 739 2 D85961
44 43 36.4 987 2 AI2011
45 42.5 36.0 101 2 T11565

probable tenascin
dihydroorotate oxi
citrate synthase-1
GP120 V3 LOOP prot
conserved hypothet
phosphoglycerate m
phosphoglycerate m
branched-chain ami
hypothetical prote
probable tyrs prot
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
peptide synthetase
rev protein - simi

Query Match 43.2%; Score 51; DB 2; Length 403;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 TWMTLQGLLDRIQAPPS 19

DB 267 TWGETWEGALDEIRALPA 284

RESULT 2

H65070

hypothetical protein b2872 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: H65070
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; UID:97426617; PMID:9278503
A;Accession: H65070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-403 <BLAT>
A;Cross-references: UNIPROT:Q46805; GB:AE000370; GB:AE00096; NID:G2367170; PIDN:AACT75910
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: succinyl-diaminopimelate desuccinylase

H65070

hypothetical protein b2872 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: H65070
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; UID:97426617; PMID:9278503
A;Accession: H65070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-403 <BLAT>
A;Cross-references: UNIPROT:Q46805; GB:AE000370; GB:AE00096; NID:G2367170; PIDN:AACT75910
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: succinyl-diaminopimelate desuccinylase

H65070

hypothetical protein b2872 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: H65070
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; UID:97426617; PMID:9278503
A;Accession: H65070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-403 <BLAT>
A;Cross-references: UNIPROT:Q46805; GB:AE000370; GB:AE00096; NID:G2367170; PIDN:AACT75910
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: succinyl-diaminopimelate desuccinylase

H65070

hypothetical protein b2872 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: H65070
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; UID:97426617; PMID:9278503
A;Accession: H65070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-403 <BLAT>
A;Cross-references: UNIPROT:Q46805; GB:AE000370; GB:AE00096; NID:G2367170; PIDN:AACT75910
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: succinyl-diaminopimelate desuccinylase

H65070

hypothetical protein b2872 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: H65070
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; UID:97426617; PMID:9278503
A;Accession: H65070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-403 <BLAT>
A;Cross-references: UNIPROT:Q46805; GB:AE000370; GB:AE00096; NID:G2367170; PIDN:AACT75910
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: succinyl-diaminopimelate desuccinylase

F;252/Active site: Cys #status predicted

Query Match 40.7%; Score 48; DB 1; Length 2225;
 Best Local Similarity 42.1%; Pred. No. 62;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 QY 4 MKTQLGLLDRIQAFPPSPH 22
 DB 440 VQTSQGLADKVKYFLPIPTH 458

RESULT 6
 S61504
 glycogen(starch) synthase (EC 2.4.1.11) isoform I precursor - garden pea
 N;Alternate names: glycogen(starch) synthase
 C;Species: Pisum sativum (garden pea)
 C;Date: 20-Jul-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
 C;Accession: S61504; S72372
 R;Dry, I.; Smith, A.; Edwards, A.; Bhattacharyya, M.; Dunn, P.; Martin, C.
 Plant J. 2, 193-202, 1992
 A;Title: Characterization of cDNAs encoding two isoforms of granule-bound starch synthase
 A;Reference number: S61504; MUID:93251108; PMID:1302049
 A;Accession: S61504
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-603 <DRY>
 A;Cross-references: UNIPROT:Q43092; EMBL:X88789; NID:g887570; PIDN:CAR61268.1; PID:g88757

A;Accession: S72372
 A;Molecule type: protein
 A;Residues: 76-77, 'X', 79, 81-88 <DRW>
 C;Superfamily: starch synthase
 C;Keywords: glycosyltransferase
 F:1-75/Domain: signal sequence #status predicted <SIG>
 F:76-603/Product: glycogen (starch) synthase isoform I #status experimental <MAT>

Query Match 39.8%; Score 47; DB 2; Length 603;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 3 WMKTLQGLLDRIQAFPPSPH 22
 DB 301 WMKA--GILESDQVFTVSPH 318

RESULT 7
 B81690
 probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqrB chain TC
 C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: B81690
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: B81690
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-503 <TET>
 A;Cross-references: UNIPROT:Q9PKB6; GB:AE002323; GB:AE002160; NID:G7190585; PIDN:AAF3938
 A;Experimental source: strain Nigg (MoPn)

C;Genetics:
 A;Gene: TC0550
 C;Keywords: oxidoreductase
 Query Match 39.0%; Score 46; DB 2; Length 503;
 Best Local Similarity 37.0%; Pred. No. 25;
 Matches 10; Conservative 6; Mismatches 5; Indels 6; Gaps 1;

QY 1 ATWMKTLQGL-----LDRIQAFPPSP 21
 DB 289 STWAESPFGTLVDQLSLDQLQNFVTS 315

RESULT 8

H86828
 fructuronate reductase (EC 1.1.1.57) [imported] - Lactococcus lactis subsp. lactis (str
 C;Species: Lactococcus lactis subsp. lactis
 C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C;Accession: H86828
 R;Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrh
 Genome Res. 11, 731-753, 2001
 A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: H86828
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-544 <STO>
 A;Cross-references: UNIPROT:Q9CF49; GB:AE005176; PID:g12724641; PIDN:AAK05730.1; GSPDB:
 A;Experimental source: strain IL1403
 C;Genetics:

A;Gene: uxuB
 C;Superfamily: mannitol 2-dehydrogenase
 C;Keywords: oxidoreductase

Query Match 39.0%; Score 46; DB 2; Length 544;
 Best Local Similarity 38.1%; Pred. No. 28;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATWMKTLQGLLDRIQAFPPSP 21
 DB 445 AAWCRYLIGINDELESFSPSP 465

RESULT 9

A35655
 peptidyl-dipeptidase A (EC 3.4.15.1), testis - mouse
 N;Alternate names: peptidyl-dipeptidase I, testis
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: A35655
 R;Howard, T.E.; Shai, S.Y.; Langford, K.G.; Martin, B.M.; Bernstein, K.E.
 Mol. Cell. Biol. 10, 4294-4302, 1990

A;Title: Transcription of testicular angiotensin-converting enzyme (ACE) is initiated w
 A;Reference number: A35655; MUID:90318396; PMID:2164636
 A;Accession: A35655
 A;Status: preliminary
 A;Molecule type: mRNA

A;Residues: 1-732 <HOW>
 A;Cross-references: UNIPROT:P22967; GB:M55333; NID:g191589; PIDN:AAA37149.1; PID:g19159
 C;Superfamily: mammalian peptidyl-dipeptidase A
 C;Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein; z

Query Match 39.0%; Score 46; DB 1; Length 732;
 Best Local Similarity 35.0%; Pred. No. 38;
 Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 WMKTLQGLLDRIQAFPPSPH 22
 DB 309 WAQTWSNIYDLVAPFPSPAPN 328

RESULT 10

AD2572
 hypothetical protein all9003 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120e
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AD2572

R;Kano, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium An
 A;Reference number: ABL1807; MUID:21595285; PMID:11759840

A;Accession: AD2572
 A;Status: preliminary

A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-1097 <KUR>
 A:Cross-references: UNIPROT:Q8YUW0; GB:AP003605; PIDN:BAW7489.1; PID:g17134934; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all19003
 A:Genome: plasmid

Query Match 39.0%; Score 46; DB 2; Length 1097;
 Best Local Similarity 38.5%; Pred. NO. 60;
 Matches 10; Conservative 5; Mismatches 5; Indels 6; Gaps 1;

QY 2 TWMTLQGLLDRIQA-----FPSPSP 21
 DB 655 SWYGTIQAILDRIQAALIFRKTFKINP 680

RESULT 11
 T47523
 DNA-binding protein-like - Arabidopsis thaliana
 N:Alternate names: protein F16L2.40
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C:Accession: T47523
 R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseid, M.; Mewes, H.W.; Rudd, S.;
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: 234468
 A:Accession: T47523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1298 <JOR>
 A:Cross-references: UNIPROT:Q9LZU7; EMBL:AL162459
 A:Experimental source: cultivar Columbia; BAC clone F16L2
 C:Genetics:
 A:Map position: 3
 A>Note: F16L2.40

Query Match 39.0%; Score 46; DB 2; Length 1298;
 Best Local Similarity 42.9%; Pred. NO. 72;
 Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATWMKTLOGLLDRIQAEPSSP 21
 DB 983 ANWLKGTQETLQIGSLPEPP 1003

RESULT 12
 A34171
 peptidyl-dipeptidase A (EC 3.4.15.1) precursor - mouse
 N:Alternate names: ACE; angiotensin-converting enzyme; carboxypeptidase; dipeptidyl car
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A34171; A29220; A61477
 R:Bernstein, K.E.; Martin, B.M.; Edwards, A.S.; Bernstein, E.A.
 J. Biol. Chem. 264, 11945-11951, 1989
 A:Title: Mouse angiotensin-converting enzyme is a protein composed of two homologous dom
 A:Reference number: A34171; MUID:83308599; PMID:2545691
 A:Accession: A34171
 A:Molecule type: mRNA
 A:Residues: 1-1312 <BER>
 A:Cross-references: UNIPROT:P09470; GB:J04947
 R:Bernstein, K.E.; Martin, B.M.; Bernstein, E.A.; Linton, J.; Striker, L.; Striker, G.
 J. Biol. Chem. 263, 11021-11024, 1988
 A:Title: The isolation of angiotensin-converting enzyme cDNA.
 A:Reference number: A29220; MUID:88298730; PMID:2841312
 A:Accession: A29220
 A:Molecule type: mRNA
 A:Residues: 1-332 <BE2>
 A:Cross-references: GB:J03940; NID:g191583; PIDN:AAA37146.1; PID:g191584
 R:Bernstein, K.E.; Martin, B.M.; Striker, L.; Striker, G.
 Kidney Int. 33, 652-655, 1988
 A:Title: Partial protein sequence of mouse and bovine kidney angiotensin converting enzy
 A:Reference number: A61477; MUID:88215372; PMID:2835538

A:Accession: A61477
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 35-54 <BE3>
 A:Experimental source: kidney
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; blood pressure control; membrane protein; peptidylaipe
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-1312/Product: peptidyl dipeptidase I #status predicted <MAT>

Query Match 39.0%; Score 46; DB 1; Length 1312;
 Best Local Similarity 35.0%; Pred. NO. 73;
 Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 WNKTLQGLLDRIQAEPSSPH 22
 DB 889 WAQTWSNIYDLVAPPPAPN 908

RESULT 13
 B83605
 Probable acetylcholinesterase aminohydrolase PA0321 [imported] - Pseudomonas aeruginosa (str
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: B83605
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: B83605
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-344 <STO>
 A:Cross-references: UNIPROT:Q916H0; GB:AE004470; GB:AE004091; NID:g9946164; PIDN:AAG037
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0321

Query Match 38.1%; Score 45; DB 2; Length 344;
 Best Local Similarity 47.1%; Pred. NO. 24;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATWMKTLOGLLDRIQA 17
 DB 256 ATWSQALQAIRIQAY 272

RESULT 14
 T49157
 citrate synthase-like protein - Arabidopsis thaliana
 N:Alternate names: protein T20N10.90
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T49157
 R:D'Angelo, M.; Verzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.;
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z25017
 A:Accession: T49157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-480 <DAN>
 A:Cross-references: UNIPROT:Q9LXS7; EMBL:AL353032; GSPDB:GX00061; ATSP:T20N10.90
 A:Experimental source: cultivar Columbia; BAC clone T20N10
 C:Genetics:
 A:Gene: ATSP:T20N10.90
 A:Map position: 3
 A:Introns: 79/3; 140/1; 168/3; 203/3; 224/3; 261/2; 301/2; 351/2; 391/3; 424/2; 461/3
 C:Superfamily: citrate synthase

Query Match 38.1%; Score 45; DB 2; Length 480;
 Best Local Similarity 53.3%; Pred. NO. 35;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 QGLLDRIQAPSPH 22
 ||:||||:|
 Db 165 QGVLDMIQSPNDVH 179

RESULT 15

S53007
 citrate synthase - cucurbit
 C:Species: Cucurbita sp. (cucurbit)
 C:Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 04-Apr-2004
 C:Accession: S53007
 R:Kato, A.; Hayashi, M.; Mori, H.; Nishimura, M.
 Plant Mol. Biol. 27, 377-390, 1995
 A:Title: Molecular characterization of a glyoxysomal citrate synthase that is synthesized
 A:Reference number: S53007; MUID:95195164; PMID:7888626
 A:Accession: S53007
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-516 <KAT>
 A:Cross-references: GB:D38132; NID:g975632; PIDN:BAA07328.1; PID:g975633
 C:Superfamily: citrate synthase

Query Match 38.1%; Score 45; DB 2; Length 516;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 QGLLDRIQAPSPH 22
 ||:||||:|
 Db 173 QGLVDIIQAMPDHAH 187

Search completed: November 10, 2004, 15:55:05
 Job time : 7.51601 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 36.8754 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-68

Perfect score: 118

Sequence: 1 ATWMKTQLGLDRIQAFSSPH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	46.6	593	2 Q7PP25	Q7pp25 anopheles g
2	53	44.9	301	2 Q73D20	Q73d20 bacillus ce
3	53	44.9	301	2 Q81YW5	Q81yw5 bacillus an
4	53	44.9	301	2 AAS39505	Aas39505 bacillus
5	53	44.9	301	2 AAT29609	Aat29609 bacillus
6	53	44.9	486	2 Q7ZWR5	Q7zwr5 xenopus lae
7	53	44.9	583	2 Q6DFB3	Q6dfb3 xenopus lae
8	52	44.1	817	1 Y1M1 SCHPO	C36033 schizosacch
9	51	43.2	201	2 Q8ALF2	Q8ali2 human immun
10	51	43.2	403	1 YGEY ECOLI	C46805 escherichia
11	51	43.2	405	2 Q6LQW1	C6lqn1 photobacter
12	51	43.2	405	2 CAG20395	Cag20395 photobact
13	50	42.4	120	2 Q9NUM7	Q9num7 homo sapien
14	50	42.4	153	2 Q8CV30	Q8cv30 mus musculu
15	50	42.4	390	2 Q95888	Q95888 homo sapien
16	50	42.4	390	2 Q91X76	Q91x76 mus musculu
17	50	42.4	461	2 Q96C80	Q96c80 homo sapien
18	50	42.4	464	2 Q9P2T2	Q9p2t2 homo sapien
19	50	42.4	473	2 Q9P2T3	Q9p2t3 homo sapien
20	50	42.4	486	2 Q9H9Z8	Q9h9z8 homo sapien
21	50	42.4	520	2 Q9H857	Q9h857 homo sapien
22	50	42.4	548	2 Q86UY8	Q86uy8 homo sapien
23	50	42.4	553	2 Q6QON3	Q6qon3 rattus norv
24	50	42.4	553	2 AAS75314	Aas75314 rattus no
25	50	42.4	4625	1 DY1A CHLRE	Q9smh3 chlamydomon
26	49.5	41.9	87	2 Q90BM5	Q90bm5 human immun
27	49.5	41.9	90	2 Q9QC08	Q9qcs8 human immun
28	49.5	41.9	352	2 Q837D2	Q837d2 enterococcu
29	49	41.5	296	2 Q9A6W1	Q9a6w1 caulobacter
30	49	41.5	301	2 Q6HNU1	Q6hnu1 bacillus th
31	49	41.5	571	2 Q81ES3	Q81es3 bacillus ce

```

32 48 40.7 213 2 QASL1 Q9as11 oryza sativ
33 48 40.7 500 2 Q6EUF8 Q6euf8 oryza sativ
34 48 40.7 2225 1 PYR1 HUMAN P27708 homo sapien
35 48 40.7 2225 1 PYR1 MESAU P08955 mesocricetu
36 48 40.7 2225 2 AAH65510 Aah65510 homo sapi
37 47.5 40.3 87 2 Q90BM2 Q90bm2 human immun
38 47.5 40.3 208 2 Q7SUP8 Q7sup8 human immun
39 47 39.8 111 2 Q9QS63 Q9qs63 human immun
40 47 39.8 112 2 Q9QS71 Q9qs71 human immun
41 47 39.8 359 2 Q9VQX3 Q9vgx3 drosophila
42 47 39.8 449 2 Q84UE8 Q84ue8 lotus japon
43 47 39.8 603 1 SSG1 PEA Q43092 pisum sativ
44 47 39.8 868 2 Q802Z5 Q802z5 human immun
45 47 39.8 1594 2 Q6C7K0 Q6c7k0 yarrowia li

```

ALIGNMENTS

RESULT 1

```

Q7PP25 ID Q7PP25 PRELIMINARY; PRT; 593 AA.
AC Q7PP25;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE ENSANGP0000010355 (Fragment).
GN Name=ENSANGG0000007866;
OS Anopheles gambiae str. FEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
RX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008900; EAA09500.2; -.
DR GO; GO:0015021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008810; F:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00850; MFS; 1.
FT NON_TER 1
FT NON_TER 593
SQ SEQUENCE 593 AA; 65943 MW; 620395C348C48F6E CRC64;

```

```

Query Match 46.6%; Score 55; DB 2; Length 593;
Best Local Similarity 52.9%; Pred. No. 9.5;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ATWMKTQLGLDRIQAF 17
Db :|||||:|:|:|
437 STWIKLEGIIMSIQTF 453

```

RESULT 2

```

Q73D20 ID Q73D20 PRELIMINARY; PRT; 301 AA.
AC Q73D20;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE Cell division inhibitor-like protein.
GN OrderedLocustNames=BCE0570;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
RX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.

```

DR PubMed=14960714;
KW Raiko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
SQ Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.,
RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic
adaptations and a large plasmid related to *Bacillus anthracis* pXOI.",
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AEO17266; AAS39505.1; -
DR TIGR; BCE0570; -
DR InterPro; IPR010099; Cons_hypoth_yfch.
DR TIGRFAMs; TIGR01777; yfch.1.
KW Cell division; Complete proteome.
SQ SEQUENCE 301 AA; 33595 MW; 3E24F59EC43DFECB CRC64;
Query Match 44.9%; Score 53; DB 2; Length 301;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 4 MKTQGLLDRIQAPSSPH 22
DB 88 IOTTGKLIKQLQALPKPH 106
RESULT 3
Q81YWS PRELIMINARY; PRT; 301 AA.
AC Q81YWS; Q813R4; Q6KXN8;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2004 (TREMELrel. 28, Last annotation update)
DE Cell division inhibitor-like protein.
GN OrderedLocusNames=BA0515, BAS0486; ORFNames=GBAA0515;
OS *Bacillus anthracis*.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.B., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Oekstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.N., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of *Bacillus anthracis* Ames and comparison to
closely related bacteria."
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RA "Bacillus anthracis comparative genomics."
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Stearns;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17025; RAP24537.1; -
DR EMBL; AEO17334; AAT29609.1; -
DR EMBL; AEO17225; AAT52817.1; -
DR TIGR; BA0515; -
DR InterPro; IPR010099; Cons_hypoth_yfch.

DR TIGRFAMs; TIGR01777; yfch.1.
KW Cell division.
SQ SEQUENCE 301 AA; 33595 MW; 4D2255DB55A224A3 CRC64;
Query Match 44.9%; Score 53; DB 2; Length 301;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 4 MKTQGLLDRIQAPSSPH 22
DB 88 IOTTGKLIKQLQALPKPH 106
RESULT 4
AAS39505 PRELIMINARY; PRT; 301 AA.
AC AAS39505;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 04-MAY-2004 (TREMELrel. 27, Last annotation update)
DE Cell division inhibitor-like protein.
GN BCE0570.
OS *Bacillus cereus* (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.,
RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic
adaptations and a large plasmid related to *Bacillus anthracis* pXOI.",
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AEO17266; AAS39505.1; -
DR TIGR; BCE0570; -
KW Cell division.
SQ SEQUENCE 301 AA; 33584 MW; 3E24F59EC43DFECB CRC64;
Query Match 44.9%; Score 53; DB 2; Length 301;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 4 MKTQGLLDRIQAPSSPH 22
DB 88 IOTTGKLIKQLQALPKPH 106
RESULT 5
AAT29609 PRELIMINARY; PRT; 301 AA.
AC AAT29609;
DT 01-JUN-2004 (TREMELrel. 27, Created)
DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
DE Cell division inhibitor-like protein.
GN GBAA0515.
OS *Bacillus anthracis* str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17334; AAT29609.1; -
DR EMBL; AEO17334; AAT29609.1; -
DR EMBL; AEO17225; AAT52817.1; -
DR TIGR; BA0515; -
DR InterPro; IPR010099; Cons_hypoth_yfch.

Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 MKTQLGLDRIQAFFSPH 22
::|||:::|||::||
Db 88 IOTKGLIKQLQALPAKH 106

RESULT 6
O7ZWR5 Q7ZWR5 PRELIMINARY; PRT; 486 AA.
ID Q7ZWR5 AC
Q7ZWR5
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2004 (TREMBLrel. 26, Last annotation update)
DE MGCS3840 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenoportinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
FAhey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917;
RX Klein S.L., Strausberg R.L., Wagner L

OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972; PubMed=11859360; DOI=10.1038/nature724;
 RX MEDLINE=21848401; Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hicalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones D., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Oliver K., O'Neill S., Saunders D., Seeger K., Sharp S.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vancstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong I., Potashkin J.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe.";
 RL Nature 415:871-880(2002).
 CC -!- SIMILARITY: SOME TO C-ELGANS T20B12.1 AND YEAST YNL313C.
 CC -!- SIMILARITY: Contains 4 TPR repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z98980; CAB11723.1; .
 DR EMBL; AL390814; CAC00549.1; .
 DR GenDB SPombe; SPAC19B12.01; .
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00515; TPR; 2.
 DR PROSITE; PS50005; TPR; 3.
 DR PROSITE; PS50293; TPR REGION; 1.
 KW Hypothetical protein; Repeat; TPR repeat.
 FT REPEAT 459 492 TPR 1.
 FT REPEAT 521 554 TPR 2.
 FT REPEAT 555 588 TPR 3.
 FT REPEAT 625 658 TPR 4.
 FT REPEAT 817 849 TPR 5.
 SQ SEQUENCE 817 AA; 94268 MW; 8B05FA55E06B6E21 CRC64;
 Query Match 44.1%; Score 52; DB 1; Length 817;
 Best Local Similarity 47.1%; Pred. No. 39;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATWMTLQGLLDRIQAF 17
 DB 747 ATWMTLQGLLDRIQAF 763
 RESULT 9
 Q8ALI2 PRELIMINARY; PRT; 201 AA.
 OS Schizosaccharomycetes pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972; PubMed=11859360; DOI=10.1038/nature724;
 RX MEDLINE=21848401; Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hicalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones D., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Oliver K., O'Neill S., Saunders D., Seeger K., Sharp S.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vancstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong I., Potashkin J.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe.";
 RL Nature 415:871-880(2002).
 CC -!- SIMILARITY: SOME TO C-ELGANS T20B12.1 AND YEAST YNL313C.
 CC -!- SIMILARITY: Contains 4 TPR repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z98980; CAB11723.1; .
 DR EMBL; AL390814; CAC00549.1; .
 DR GenDB SPombe; SPAC19B12.01; .
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00515; TPR; 2.
 DR PROSITE; PS50005; TPR; 3.
 DR PROSITE; PS50293; TPR REGION; 1.
 KW Hypothetical protein; Repeat; TPR repeat.
 FT REPEAT 459 492 TPR 1.
 FT REPEAT 521 554 TPR 2.
 FT REPEAT 555 588 TPR 3.
 FT REPEAT 625 658 TPR 4.
 FT REPEAT 817 849 TPR 5.
 SQ SEQUENCE 817 AA; 94268 MW; 8B05FA55E06B6E21 CRC64;
 Query Match 44.1%; Score 52; DB 1; Length 817;
 Best Local Similarity 47.1%; Pred. No. 39;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATWMTLQGLLDRIQAF 17
 DB 747 ATWMTLQGLLDRIQAF 763
 RESULT 9
 Q8ALI2 PRELIMINARY; PRT; 201 AA.

AC Q8ALI2;
 DT 01-WAR-2003 (TREMELrel. 23, Created)
 DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN Name=env;
 OS Human immunodeficiency virus 1.
 OX Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22440026; PubMed=12551997;
 RX Gordon M., De Oliveira T., Bishop K., Coovadia H.M., Madurai L.,
 RA Engelbrecht S., van Rensburg E.J., Mosam A., Smith A., Cassol S.;
 RA "Molecular characteristics of human immunodeficiency virus type 1
 RT subtype C viruses from KwaZulu-Natal, South Africa: implications for
 RT vaccine and antiretroviral control strategies.";
 RL J. Virol. 77:2581-2599(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Tarin M., De Oliveira T., Bishop K., Coovadia H.M., Madurai S.,
 RA Engelbrecht S., van Rensburg E.J., Mosam A., Smith A., Cassol S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV137054; RAN16003.1; .
 DR GO; GO:0013028; C:Viral capsid; IEA.
 DR GO; GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
 FT NON_TER 201
 FT NON_TER 201
 SQ SEQUENCE 201 AA; 22418 MW; 1D99A7FA507028F4 CRC64;
 Query Match 43.2%; Score 51; DB 2; Length 201;
 Best Local Similarity 47.1%; Pred. No. 14;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATWMTLQGLLDRIQAF 17
 DB 67 ANWTKTQGVSKLREY 83
 RESULT 10
 YGEY_ECOLI STANDARD; PRT; 403 AA.
 ID YGEY_ECOLI
 AC Q46805;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical protein ygeY.
 GN Name=ygeY; OrderedLocusNames=b2872, c3450, z4211, Ecs3745;
 OS Escherichia coli,
 OS Escherichia coli O6, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 217992, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1233-1238(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
 RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

DT	05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE	Putative deacetylase.
GN	O=OrderedLocusNames=PBPRA1992;
OS	Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Photobacterium.
OX	NCBI_TaxID=74109;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Vezi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA	Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA	Valle G.;
RT	"Genome analysis of Photobacterium profundum reveals the complexity of
RL	high pressure adaptations.";
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; CR378669; CAG20395.1; -.
DR	InterPro; IPR0011261; ArgE_dapE.
DR	InterPro; IPR006035; Arg_agm_form.
DR	InterPro; IPR002933; Peptidase_M20.
DR	Pfam; PF01546; Peptidase_M20_1.
DR	PROSITE; PS00758; ARG_E-DAP_E_CFG5_1; UNKNOWN_1.
DR	PROSITE; PS00148; ARGINASE_2; UNKNOWN_1.
KW	Complete proteome.
SQ	SEQUENCE 405 AA; 45299 MW; 53BB7A9CD45DA868 CRC64;
Query Match	43.2%; Score 51; DB 2; Length 405;
Best Local Similarity	50.0%; Pred.No. 27;
Matches	9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy	2 TWMKTLQGLDLRIQAPPS 19
Dd	: : :
Dd	267 TWGETWEGALDEIRALPA 284
RESULT 12	
CAG20395	
ID	CAG20395 PRELIMINARY; PET; 405 AA.
AC	CAG20395;
DT	10-MAY-2004 (TReMBLrel. 27, Created)
DT	10-MAY-2004 (TReMBLrel. 27, Last sequence update)
DT	10-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE	Putative deacetylase.
GN	PBPRA1992.
OS	Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Photobacterium.
OX	NCBI_TaxID=74109;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ss9;
RA	Vezi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA	Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA	Valle G.;
RT	"Genome Analysis of Photobacterium profundum reveals the complexity of
RL	high pressure adaptations.";
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ss9;
RA	Cestaro A.;
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; CR378669; CAG20395.1; -.
SQ	SEQUENCE 405 AA; 45299 MW; 53BB7A9CD45DA868 CRC64;
Query Match	43.2%; Score 51; DB 2; Length 405;
Best Local Similarity	50.0%; Pred.No. 27;
Matches	9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy	2 TWMKTLQGLDLRIQAPPS 19
Dd	: : :
Dd	267 TWGETWEGALDEIRALPA 284

RESULT 13

Q9CYN7 PRELIMINARY; PRT; 120 AA.
 AC Q9CYN7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein FLJ11266.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX PubMed=14702039;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Hirakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirakawa K., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Itakura S., Fukuzumi Y.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi K., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugeno J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK002128; BAA92095.1; -;
 DR InterPro; IPR008380; Purine nucl.
 DR Pfam; PF05761; 5_nucleotid; 1.
 SQ SEQUENCE 120 AA; 14303 MW; C9605BF684726B2C CRC64;

Query Match 42.4%; Score 50; DB 2; Length 120;
 Best Local Similarity 57.1%; Pred. No. 12;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWMTKQLGLDRIO 15
 ||:|||||
 DB 2 TWLQTLGLLEFQM 15
 ||:|||||

RESULT 14

Q9CY30 PRELIMINARY; PRT; 153 AA.
 AC Q9CY30
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 13 days embryo liver cDNA, RIKEN full-length enriched
 DE library, clone:2510015F01 product:cDNA FLJ13933 FIS, CLONE
 DE Y79AA1000782, WEAKLY SIMILAR TO CYTOSOLIC PURINE 5'-NUCLEOTIDASE (EC
 DE 3.1.3.5) homolog.
 GN Name=2510015F01Rik;

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RA The FANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanganaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Horii F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Sano H., Sasaki D., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Shinagawa A., Shiraki T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK010966; BAB27298.1; -;
 DR MGD; MGI:1917271; 2510015F01Rik.
 DR InterPro; IPR008380; Purine nucl.
 DR Pfam; PF05761; 5_nucleotid; 1.
 SQ SEQUENCE 153 AA; 18301 MW; D5397211877B00E7 CRC64;

Query Match 42.4%; Score 50; DB 2; Length 153;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 2 TWMTLQGLDRIQAF 17
DB 32 TWQALTGLLERMQTY 47

RESULT 15
Q95888 PRELIMINARY; PRT; 390 AA.
AC C95888;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96207227; PubMed=9619474;
RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
RT "A 'double adaptor' method for improved shotgun library
RL construction.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97264341; PubMed=91101174;
RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
RT "Large-scale concatenation cDNA sequencing.";
RN [3]
RP SEQUENCE FROM N.A.
RA Mei G., Yu W., Gibbs R.A.;
RL Submitted (FEF-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131781; AAD2004.1; -.
DR InterPro; IPR008380; Purine_nucl.
DR Pfam; PF05761; 5_nucleotid; 1.
KW Hypothetical protein.
SQ SEQUENCE 390 AA; 45998 MW; D51C366CDC4A108B CRC64;

Query Match 42.4%; Score 50; DB 2; Length 390;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 TWMTLQGLDRIQAF 17
DB 269 TWQALTGLLERMQTY 284

```

Search completed: November 10, 2004, 15:53:30
Job time : 38.9524 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 13.879 Seconds
(without alignments)
143.349 Million cell updates/1000000 cells

Title: US-10-092-750-69

Perfect score: 161
Sequence: 1 EANKQPKPNNSSTAYNFTGVSLPSYKP 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```
Database : Issued Patents AA: *
1: /cqn2 6/ptodata/1/iaa/5A COMB.rep:*
```

1: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.per.*

```
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_5/ptodata/1/1aa/6C_COMB.pep:
```

```
5: /cgn2_6/production/1/aa/backfile/need:
5: /cgn2_6/production/1/aa/backfile/need:
5: /cgn2_6/production/1/aa/backfile/need:
5: /cgn2_6/production/1/aa/backfile/need:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	50	31.1	926	1	US-08-159-340A-2	Sequence 2, Appli
2	48	29.8	717	4	US-08-924-623C-5	Sequence 5, Appli
3	47.5	29.5	2512	2	US-08-801-263A-9	Sequence 9, Appli
4	47.5	29.5	2512	3	US-09-103-248-9	Sequence 9, Appli
5	47.5	29.5	2512	4	US-09-367-764-9	Sequence 9, Appli
6	47	29.2	501	4	US-09-248-796A-18735	Sequence 18735, A
7	46.5	28.9	632	4	US-09-538-092-503	Sequence 503, App
8	46.5	28.9	682	4	US-09-101-532A-8403	Sequence 8403, Ap
9	46	28.6	84	4	US-09-252-991A-24609	Sequence 24609, A
10	46	28.6	308	4	US-09-328-352-7565	Sequence 7565, Ap
11	45.5	28.3	60	4	US-09-248-796A-26885	Sequence 26885, A
12	45.5	28.3	776	1	US-07-603-133B-20	Sequence 20, Appl
13	45.5	28.3	776	3	US-08-089-397A-15	Sequence 15, Appl
14	45	28.0	67	4	US-09-101-532A-4745	Sequence 4745, Ap
15	45	28.0	114	4	US-09-252-991A-26937	Sequence 26937, A
16	45	28.0	192	4	US-09-270-767-38337	Sequence 38337, A
17	45	28.0	192	4	US-09-270-767-53554	Sequence 53554, A
18	45	28.0	270	4	US-09-134-000C-3553	Sequence 3553, Ap
19	45	28.0	442	4	US-08-328-352-5419	Sequence 5419, Ap
20	45	28.0	620	1	US-08-325-071-65	Sequence 65, Appl
21	45	28.0	620	3	US-08-461-004A-65	Sequence 65, Appl
22	45	28.0	650	1	US-08-325-071-56	Sequence 56, Appl
23	45	28.0	650	1	US-08-325-071-59	Sequence 59, Appl
24	45	28.0	650	1	US-08-325-071-63	Sequence 63, Appl
25	45	28.0	650	1	US-08-325-071-67	Sequence 67, Appl
26	45	28.0	650	3	US-08-461-004A-56	Sequence 56, Appl
27	45	28.0	650	3	US-08-461-004A-59	Sequence 59, Appl

28	45	28.0	650	3	US-08-461-0048-63	Sequence 63, Appl
29	45	28.0	650	3	US-08-461-0048-67	Sequence 67, Appl
30	45	28.0	688	1	US-08-461-0041-57	Sequence 57, Appl
31	45	28.0	688	3	US-08-461-0041-57	Sequence 57, Appl
32	44.5	27.6	761	4	US-08-461-0048-35	Sequence 35, Appl
33	44	27.3	69	4	US-09-138-4524-335	Sequence 335, Appl
34	44	27.3	69	4	US-09-270-767-34348	Sequence 34348, A
35	44	27.3	69	4	US-09-270-767-34348	Sequence 34348, A
36	44	27.3	76	4	US-09-248-796A-23262	Sequence 23262, A
37	44	27.3	87	4	US-09-248-796A-24444	Sequence 24444, A
38	44	27.3	111	4	US-09-513-599C-7348	Sequence 7348, Ap
39	44	27.3	134	4	US-09-248-796A-27538	Sequence 27538, A
40	44	27.3	140	4	US-09-282-991A-25306	Sequence 2306, A
41	44	27.3	380	2	US-09-248-796A-19662	Sequence 19662, A
42	44	27.3	462	2	US-08-898-976-2	Sequence 2, Appl
43	44	27.3	463	2	US-08-898-976-4	Sequence 4, Appl
44	44	27.3	463	4	US-09-252-991A-27686	Sequence 27686, A
45	44	27.3	499	4	US-08-594-193-4	Sequence 4, Appl
46	44	27.3	532	3	US-09-008-271A-12	Sequence 12, Appl

ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match      31.1%; Score 50; DB 1; Length 926;
Best Local Similarity 34.5%; Pred. No. 65;
Matches 10; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY      2 ANRQKPKNNSTAYYNFTGVSILPSYKP 30
          :|:::|:::|:::|
Db      410 SNTISNPKNNSNLYSIGLSLSSGSPSP 438

```

```

RESULT 2
US-08-924-629C-5
; Sequence 5, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 717
; TYPE: PRT
; ORGANISM: lcaC;
US-08-924-629C-5

Query Match 29.8%; Score 48; DB 4; Length 717;
Best Local Similarity 44.4%; Pred. No. 95;
Matches 12; Conservative 5; Mismatches 4; Indels 6; Gaps 2;

QY 10 NNSSTAYYN--FTGVSI-----PSYKP 30
DB 120 NMTKAFNEEWTGVSIAPNTYKP 146

RESULT 3
US-08-801-263A-9
; Sequence 9, Application US/08801263A
; Patent No. 5811407
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bell Seltzer Park & Gibson, P.A.
; CITY: Charlotte
; STATE: No. 5811407th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,263A
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175

US-08-924-629C-5
; Sequence 5, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 717
; TYPE: PRT
; ORGANISM: lcaC;
US-08-924-629C-5

Query Match 29.8%; Score 48; DB 4; Length 717;
Best Local Similarity 44.4%; Pred. No. 95;
Matches 12; Conservative 5; Mismatches 4; Indels 6; Gaps 2;

QY 10 NNSSTAYYN--FTGVSI-----PSYKP 30
DB 120 NMTKAFNEEWTGVSIAPNTYKP 146

RESULT 3
US-08-801-263A-9
; Sequence 9, Application US/08801263A
; Patent No. 5811407
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bell Seltzer Park & Gibson, P.A.
; CITY: Charlotte
; STATE: No. 5811407th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,263A
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175

```

```

; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2512 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-801-263A-9

Query Match 29.5%; Score 47.5; DB 2; Length 2512;
Best Local Similarity 42.3%; Pred. No. 5.3e-02;
Matches 11; Conservative 6; Mismatches 6; Indels 3; Gaps 1;

QY 2 ANRKQPKP---NNSSTAYYNFTGVSI 24
DB 1782 AARKEPTPASNSSESLSLHLSFGVSM 1807

RESULT 4
US-09-102-248-9
; Sequence 9, Application US/09102248
; Patent No. 6008035
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bell Seltzer Park & Gibson, P.A.
; CITY: Charlotte
; STATE: No. 6008035th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,248
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/801,263
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2512 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-102-248-9

Query Match 29.5%; Score 47.5; DB 3; Length 2512;
Best Local Similarity 42.3%; Pred. No. 5.3e-02;
Matches 11; Conservative 6; Mismatches 6; Indels 3; Gaps 1;

QY 2 ANRKQPKP---NNSSTAYYNFTGVSI 24
DB 1782 AARKEPTPASNSSESLSLHLSFGVSM 1807

RESULT 5
US-09-367-764-9

```


; Sequence 9, Application US/09367764
; Patent No. 6583121
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6583121th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09367,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/801,263
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2512 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-367-764-9

Query Match 29.5%; Score 47.5; DB 4; Length 2512;
Best Local Similarity 42.3%; Pred. No. 5.3e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 3; Gaps 1;

QY 2 ANRKQPKP--NNSSTAYNFTGVSI 24
| | | | | : : : : :
Db 1782 AARKEPTPPASNSSESLHLSFGVSM 1807

RESULT 6
US-09-248-796A-18735
; Sequence 18735, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18735
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18735

Query Match 29.2%; Score 47; DB 4; Length 501;
Best Local Similarity 50.0%; Pred. No. 86;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 NNSSTAYNFTGVSIPLPS 27
| | | | | : : : : :
Db 474 NNNSSSYNVTNHSQSPS 491

RESULT 7
US-09-538-092-503
; Sequence 503, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 503
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YKL204W
US-09-538-092-503

Query Match 28.9%; Score 46.5; DB 4; Length 632;
Best Local Similarity 34.5%; Pred. No. 1.4e-02;
Matches 10; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 3 NRKQPKPNSSTAYNFTGVS-ILPSYKP 30
| | | | | : : : : :
Db 71 NKESNKKKNTAMLNFSLSLTDPTNYP 99

RESULT 8
US-09-107-532A-6403
; Sequence 6403, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998

```

; APPLICATION NUMBER: 60/051571
; FILING DATE: JULY 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneko
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6403:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURES:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...682
; SEQUENCE DESCRIPTION: SEQ ID NO: 6403:
US-09-107-532A-6403

```

```

Query Match      28.8%; Score 46.5; DB 4; Length 682;
Best Local Similarity 40.9%; Pred. No. 1.5e+02;
Matches 9; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

```

```

QY      7 PKPNNSTAYNFTGVSLPSY 28
      |||||: : ||||: : :
Db      252 PKPNNNYS-GHGTGVAPTF 272

```

RESULT 9

```

US-09-252-991A-24609
; Sequence 24609, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24609
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24609

```

```

Query Match      28.6%; Score 46; DB 4; Length 84;
Best Local Similarity 37.5%; Pred. No. 13;
Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

```

```

QY      7 PKPNNSTAYNFTGVSLPSYKP 30
      ||: : ||: : ||: : ||
Db      1 PRAGWTSNLRHFNFGVNAWPSIKP 24

```

RESULT 10

```

US-09-328-352-7565
; Sequence 7565, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA

```

```

; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7565
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7565

```

```

Query Match      28.6%; Score 46; DB 4; Length 308;
Best Local Similarity 42.9%; Pred. No. 67;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      7 PKPNNSTAYNFTGVSLPS 27
      |||||: : ||||: : :
Db      203 PKPNNSTTIQALFTKGLVPA 223

```

RESULT 11

```

US-09-248-796A-26885
; Sequence 26885, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26885
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26885

```

```

Query Match      28.3%; Score 45.5; DB 4; Length 60;
Best Local Similarity 45.5%; Pred. No. 10;
Matches 10; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

```

```

QY      2 ANRKQPKPNNSTAYNFTGVS 23
      |||||: : ||||: : :
Db      29 ANR-VAKSGNSTFSYFNFAIS 49

```

RESULT 12

```

US-07-603-133B-20
; Sequence 20, Application US/07603133B
; Patent No. 5298244
; GENERAL INFORMATION:
; APPLICANT: Redmond, Mark J.
; APPLICANT: Ijaz, Mohammed K.
; APPLICANT: Parker, Michael D.
; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR
; TITLE OF INVENTION: USE IN A VACCINE TO ROTAVIRAL DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

MOLECULE IDENTIFICATION
US-08-089-397A-15

RESULT 15
US-09-252-991A-26937
; Sequence 26937, Application US/09252991A
: Patent No. 6551795

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26937
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-26937

Query Match      28.0%; Score 45; DB 4; Length 114;
Best Local Similarity 42.1%; Pred. No. 28;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      4 RKQPKPNNSSTAYNFTGV 22
DB      96 RKQPKPNKLSNISRFAGI 114

```

Search completed: November 10, 2004, 15:57:21
Job time : 13.879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 72.4911 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-69

Perfect score: 161
Sequence: 1 EANKQKPKNNSTAYNFTGVSIPLSYKP 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161	100.0	30	14	US-10-092-750-69
2	59	36.6	771	17	US-10-739-930-5692
3	55	34.2	423	14	US-10-156-761-12135
4	54.5	33.9	65	15	US-10-424-599-231460
5	53	32.9	40	15	US-10-424-599-283698
6	52.5	32.6	455	15	US-10-424-599-237197
7	52	32.3	68	15	US-10-424-599-192571
8	52	32.3	130	17	US-10-425-115-190072
9	52	32.3	211	16	US-10-437-963-123010
10	51	31.7	230	14	US-10-369-493-2545
11	51	31.7	268	17	US-10-425-115-230249
12	50	31.1	360	17	US-10-425-115-217690
13	50	31.1	381	15	US-10-425-114-59044

14	50	31.1	502	15	US-10-362-010-27	Sequence 27, Appl
15	50	31.1	926	14	US-10-369-493-1623	Sequence 1623, Ap
16	49.5	30.7	201	15	US-10-424-599-253829	Sequence 253829,
17	49	30.4	231	15	US-10-335-977-5010	Sequence 5010, Ap
18	49	30.4	258	15	US-10-425-114-43900	Sequence 43900, A
19	49	30.4	376	15	US-10-335-977-5011	Sequence 5011, Ap
20	49	30.4	440	15	US-10-424-599-272561	Sequence 272561,
21	49	30.4	446	15	US-10-425-114-49457	Sequence 49457, A
22	49	30.4	458	16	US-10-437-963-180243	Sequence 180243,
23	49	30.4	461	15	US-10-425-114-56187	Sequence 56187, A
24	49	30.4	482	15	US-10-424-599-247789	Sequence 247789,
25	49	30.4	734	10	US-09-882-227-346	Sequence 346, App
26	49	30.4	842	15	US-10-282-122A-58988	Sequence 58988, A
27	49	30.4	1036	16	US-10-322-281-160	Sequence 160, App
28	49	30.4	1254	16	US-10-437-963-109926	Sequence 109926,
29	49	30.4	1781	9	US-09-995-749A-2	Sequence 2, Appli
30	48.5	30.1	361	15	US-10-282-122A-61259	Sequence 61259, A
31	48	29.8	23	10	US-09-759-130B-295	Sequence 295, App
32	48	29.8	23	14	US-10-188-123-25	Sequence 25, Appl
33	48	29.8	23	14	US-10-188-495-25	Sequence 25, Appl
34	48	29.8	23	16	US-10-741-790-295	Sequence 295, App
35	48	29.8	57	17	US-10-425-115-348861	Sequence 348861,
36	48	29.8	174	17	US-10-425-115-227353	Sequence 227353,
37	48	29.8	353	10	US-09-759-130B-283	Sequence 283, App
38	48	29.8	353	14	US-10-188-123-13	Sequence 13, Appl
39	48	29.8	353	14	US-10-188-495-13	Sequence 13, Appl
40	48	29.8	353	16	US-10-741-790-283	Sequence 283, App
41	48	29.8	371	10	US-09-759-130B-281	Sequence 281, App
42	48	29.8	371	14	US-10-188-123-11	Sequence 11, Appl
43	48	29.8	371	14	US-10-188-495-11	Sequence 11, Appl
44	48	29.8	371	16	US-10-741-790-281	Sequence 281, App
45	48	29.8	373	9	US-09-820-893-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-10-092-750-69
; Sequence 69, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Hammon, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-69

Query Match 100.0%; Score 161; DB 14; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.8e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EANKQKPKNNSTAYNFTGVSIPLSYKP 30
Db 1 EANKQKPKNNSTAYNFTGVSIPLSYKP 30

RESULT 2
US-10-739-930-5692
; Sequence 5692, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377) B
; CURRENT APPLICATION NUMBER: US/10/739.930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5692
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(771)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C122594_1.p
US-10-739-930-5692

Query Match 36.6%; Score 59; DB 17; Length 771;
Best Local Similarity 41.7%; Pred. No. 11;
Matches 15; Conservative 5; Mismatches 6; Indels 10; Gaps 2;

QY 1 EARRKQPK-----PNNSTAYNFTGVSLPSYKP 30
DB 477 EQKPTPKQIAVIPNNASTA-----TAAALLPSHKP 508

RESULT 3
US-10-156-761-12135
; Sequence 12135, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12135
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12135

Query Match 34.2%; Score 55; DB 14; Length 423;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 14; Conservative 4; Mismatches 8; Indels 8; Gaps 2;

QY 4 RKQPKPNNSTAYNFTGVSLPSYKP-----YKP 30
DB 102 RKVAKANNASMTYF-LSGVYLLPEEKRDLYKP 132

RESULT 4
US-10-424-599-231460
; Sequence 231460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 231460
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_5102C.1.p
US-10-424-599-231460

Query Match 33.9%; Score 54.5; DB 15; Length 65;
Best Local Similarity 50.0%; Pred. No. 2.8;
Matches 11; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 9 PNNSTAYNFTGVSLPSYKP 30
DB 44 PSNVSTAHYAFS-VQFIPGYTP 64

RESULT 5
US-10-424-599-283698
; Sequence 283698, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283698
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98202C.1.p
US-10-424-599-283698

Query Match 32.9%; Score 53; DB 15; Length 40;
Best Local Similarity 43.5%; Pred. No. 2.7;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 NRKQPKPNNSTAYNFTGVSL 25
DB 12 NRRQHKDKESYGRYYSFTKITLL 34

RESULT 6
US-10-424-599-237197
; Sequence 237197, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 237197
; LENGTH: 455
; TYPE: PRT

```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56213C.1.pap
US-10-424-599-237197

Query Match      32.6%; Score 52.5; DB 15; Length 455;
Best Local Similarity 41.9%; Pred. No. 55;
Matches 13; Conservative 6; Mismatches 3; Indels 9; Gaps 2;

QY  9 PNNSSST-----AYNFTGVSI---LPSYKP 30
    ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
Db  16 PNTNSTPSVNSAPSYTGVSSVPPPIPTYP 46

RESULT 7
US-10-424-599-192571
; Sequence 192571, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 192571
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(68)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_15913C.1.pap
US-10-424-599-192571

Query Match      32.3%; Score 52; DB 15; Length 68;
Best Local Similarity 43.8%; Pred. No. 7;
Matches 14; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY  1 BANRQPKPNNSSTAYN--FTGVSIILPSYKP 30
    |||  :||:|  |||  |||  |||  |||  |||  |||  |||  |||
Db  2 BANSGASEPNSNATITINVKFSGVSIPIISIP 33

RESULT 8
US-10-425-115-190072
; Sequence 190072, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 190072
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(130)
; OTHER INFORMATION: unsure at all Xaa locations
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_104928C.1.pap
US-10-425-115-190072

Query Match      32.3%; Score 52; DB 17; Length 130;
Best Local Similarity 41.7%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY  4 RKQPKPNNSSTAYNFTGVSIILPS 27
    :  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  45 KPDPFPNQKLCAYFFPKAVDILPT 68

RESULT 9
US-10-437-963-123010
; Sequence 123010, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 123010
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(211)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_25886C.1.pap
US-10-437-963-123010

Query Match      32.3%; Score 52; DB 16; Length 211;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 12; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY  1 BANRQPKPNNSSTAYNFTGVSIILPSYKP 30
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  127 EARRAPTPPSSLPSSSPTAAKLPPSPPP 156

RESULT 10
US-10-369-493-2545
; Sequence 2545, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2545
; LENGTH: 230
```

```
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2545

Query Match      31.7%; Score 51; DB 14; Length 230;
Best Local Similarity 37.0%; Pred. No. 41;
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 4 RKQPKNNSSSTAYNFTGVSLPSYKP 30
Db 106 KKKPTPPPSDGLYYVFRGKRKKSRP 132

RESULT 11
US-10-425-115-230249
; Sequence 230249, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 230249
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; LOCATION: (1)..(268)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_141580C.1.pep
US-10-425-115-230249

Query Match      31.7%; Score 51; DB 17; Length 268;
Best Local Similarity 58.8%; Pred. No. 49;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 14 TAYNFTGVSLPSYKP 30
Db 147 TANLHTGVMLDPSYQP 163

RESULT 12
US-10-425-115-217690
; Sequence 217690, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 217690
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_130130C.1.pep
US-10-425-115-217690

Query Match      31.1%; Score 50; DB 17; Length 360;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 EANKQPKNNSSSTAYNFTGVSLP 26
Db 301 EADRRKPNNSMNGELFESITGGHLMP 326

RESULT 13
US-10-425-114-59044
; Sequence 59044, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59044
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700076533_FLI.pep
US-10-425-114-59044

Query Match      31.1%; Score 50; DB 15; Length 381;
Best Local Similarity 34.6%; Pred. No. 1e+02;
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 EANKQPKNNSSSTAYNFTGVSLP 26
Db 322 EADRRKPNNSMNGELFESITGGHLMP 347

RESULT 14
US-10-362-010-27
; Sequence 27, Application US/10362010
; Publication No. US20040038247A1
; GENERAL INFORMATION:
; APPLICANT: Brenner, Sidney
; APPLICANT: Venkatesh, Byrappa
; APPLICANT: Tan, Yin, Fwee
; TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS INCLUDING A NOVEL T-CELL ACTIVE PROMOTER
; TITLE OF INVENTION: AND PHARMACEUTICAL COMPOSITIONS AND METHODS UTILIZING SAME FOR
; TITLE OF INVENTION: REGULATING T-CELL MEDIATED IMMUNE RESPONSE
; FILE REFERENCE: 01/22004
; CURRENT APPLICATION NUMBER: US/10/362,010
; CURRENT FILING DATE: 2003-08-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Fugu rubripes
US-10-362-010-27

Query Match      31.1%; Score 50; DB 15; Length 502;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 EANKQPKNNSSSTAYNFP 19
Db 449 EQNYRMPKPNCPDGLYNF 467
```



```

RESULT 15
US-10-369-493-1623
; Sequence 1623, Application US/10369493
; Publication No. US2003033675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1623
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1623

Query Match      31.1%; Score 50; DB 14; Length 926;
Best Local Similarity 34.5%; Pred. No. 2.9e+02;
Matches 10; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy      2 ANRKQPKNNSTAYNYFTGVSYLPSYKP 30
Db      410 .SNYTSFNPKNNSNLVSISSLSISSPSP 438

Search completed: November 11, 2004, 07:41:46
Job time : 73.4911 secs

```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 10, 2004, 14:52:32 ; Search time 10.2491 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-69
Perfect score: 161
Sequence: 1 EANKQPKPNSSGTAYNFTGVSLPSYKP 30
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	36.6	771	2	145967
2	52	32.3	158	2	D89829
3	52	32.3	228	2	B72655
4	52	32.3	3014	1	JC5620
5	51.5	32.0	3262	2	A42137
6	51	31.7	230	2	S41043
7	51	31.7	230	2	T40233
8	51	31.7	864	2	JC4624
9	50.5	31.4	218	2	AG3554
10	50	31.1	926	2	S39344
11	50	31.1	1711	2	T18429
12	49	30.4	563	2	C71529
13	49	30.4	673	2	T36717
14	49	30.4	841	2	E71808
15	49	30.4	842	2	H64694
16	49	30.4	1829	2	T26135
17	48.5	30.1	354	2	AC2865
18	48.5	30.1	354	2	B97642
19	48.5	30.1	355	2	E95990
20	48.5	30.1	391	2	T14580
21	48.5	30.1	436	2	T14816
22	48.5	30.1	963	2	T40873
23	48	29.8	339	2	H90166
24	48	29.8	401	2	S71167
25	48	29.8	420	2	T32157
26	48	29.8	468	2	T48615
27	48	29.8	1711	2	T06119
28	47.5	29.5	2512	1	MMWVS
29	47	29.2	225	2	F96742

30 47 29.2 357 2 E86823
31 47 29.2 395 2 E82334
32 47 29.2 439 2 T47713
33 47 29.2 544 2 G64483
34 47 29.2 617 2 C84922
35 46.5 28.9 411 2 AF2320
36 46.5 28.9 522 2 G86259
37 46.5 28.9 632 2 S38042
38 46.5 28.9 1512 2 G85090
39 46.5 28.9 1534 2 S96004
40 46 28.6 239 2 S60083
41 46 28.6 280 2 A75526
42 46 28.6 288 2 T41101
43 46 28.6 367 2 AD1786
44 46 28.6 369 2 AF1410
45 46 28.6 437 2 I57942

ALIGNMENTS

RESULT 1
T45967
Hypothetical protein F7J8.250 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45967
R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Le
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23018
A:Accession: T45967
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-771 <BEV>
A:Cross-references: UNIPROT:Q9LFA9; EMBL:AL1137189
A:Experimental source: cultivar Columbia; BAC clone F7J8
C:Genetics:
A:Map position: 5
A:Introns: 91/3; 248/3; 319/1; 374/1; 396/2; 431/3; 462/3; 507/1; 610/3; 628/1; 672/3;
A:Note: F7J8.250

Query Match 36.6%; Score 59; DB 2; Length 771;
Best Local Similarity 41.7%; Pred. No. 2.4;
Matches 15; Conservative 5; Mismatches 6; Indels 10; Gaps 2;
QY 1 EANKQPK-----PNNSTAYNFTGVSLPSYKP 30
DB 477 EQKPTPKQIAVIPNNASTA---TAAALLPSHKP 508

RESULT 2
D89829
conserved hypothetical protein SA0559 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89829
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89829
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <KUR>
A:Cross-references: UNIPROT:Q99W10; GB:BA000018; PID:gl3700494; PIDN:BAB41791.1; GSPDB:
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0559

Query Match 32.3%; Score 52; DB 2; Length 158;
Best Local Similarity 47.6%; Pred. No. 3.9;

A;Cross-references: EMBL:Z74365; NID:gi1431526; PIDN:CXA98887.1; PID:gi1431527; MIPS:YDR06
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: SGD:DOM4; SSV7; UBP4
 A;Cross-references: SGD:S0002476; MIPS:YDR069c
 A;Map position: 4R
 C;Superfamily: deubiquinating enzyme SSV7

Query Match 31.1%; Score 50; DB 2; Length 926;
 Best Local Similarity 34.5%; Pred. No. 59;
 Matches 10; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 ANRKQPKNNSTAYNFTGVSYLPSYKP 30
 Db 410 SNYTSFNPKNNSNIYSLSSLSISSPSP 438

RESULT 11
 T18429
 hypothetical protein C0345w - malaria parasite (Plasmodium falciparum)
 C;Species: Plasmodium falciparum
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T18429
 R;Lawson, D.; Bowman, S.; Barrell, B.
 submitted to the EMBL Data Library, August 1997
 A;Reference number: Z18935
 A;Accession: T18429
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-1711 <LAW>
 A;Cross-references: UNIPROT:O77322; EMBL:Z98547; NID:e1325376; PID:e1325381; PIDN:CAB111
 C;Genetics:
 A;Note: PFC0345W

Query Match 31.1%; Score 50; DB 2; Length 1711;
 Best Local Similarity 57.9%; Pred. No. 1.2e+02;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 10 NNSSTAYNFTGVSYLPSY 28
 Db 901 NNISTYKFNFSANLIEPSY 919

RESULT 12
 C71529
 hypothetical protein CT326 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
 C;Species: Chlamydia trachomatis
 C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
 C;Accession: C71529
 R;Stephenson, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
 A;Reference number: A71570; MUID:99000809; PMID:9784136
 A;Accession: C71529
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-563 <ARN>
 A;Cross-references: UNIPROT:O84328; GB:AE001305; GB:AE001273; NID:g3328737; PIDN:AAC6791
 A;Experimental source: serotype D, strain UW-3/Cx
 A;Gene: CT326

Query Match 30.4%; Score 49; DB 2; Length 563;
 Best Local Similarity 46.2%; Pred. No. 46;
 Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 2 ANRKQPKNNSTAYNFTGVSYLPS 27
 Db 490 ANSTTSPSLNNTCTCGNSTTREVLP 515

RESULT 13
 T136717

probable serine/threonine protein kinase - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T136717
 R;Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A;Reference number: Z21612
 A;Accession: T136717
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-673 <MUR>
 A;Cross-references: UNIPROT:Q9XA16; EMBL:AL079308; PIDN:CAB45215.1; GSPDB:GN00070; SCOE
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOE:DB:SCHE9.18
 C;Superfamily: Mycobacterium tuberculosis probable pknB protein; protein kinase homolog

Query Match 30.4%; Score 49; DB 2; Length 673;
 Best Local Similarity 36.0%; Pred. No. 57;
 Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 EANKQPKNNSTAYNFTGVSYL 25
 Db 337 DRRRQPKKNTSTIFLVLAGVLV 361

RESULT 14
 E71808
 probable iron (III) dicitrate transport protein - Helicobacter pylori (strain J99)
 C;Species: Helicobacter pylori
 A;Variety: Strain J99
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat
 A;Reference number: A71800; MUID:99120557; PMID:9923682
 A;Accession: E71808
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-841 <ARN>
 A;Cross-references: UNIPROT:Q9ZJ88; GB:AE001565; GB:AE001439; NID:g4156044; PIDN:AAD070
 A;Experimental source: strain J99
 C;Genetics:
 A;Gene: feca_3
 C;Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homolo

Query Match 30.4%; Score 49; DB 2; Length 841;
 Best Local Similarity 60.0%; Pred. No. 73;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 QPKPNNSTAYNFT 20
 Db 826 QPAPGRSVTAYLNYT 840

RESULT 15
 H64694
 iron(III) dicitrate transport protein - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C;Accession: H64694
 R;Tomb, J.F.; White, C.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L
 Nature 388, 539-547, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A;Reference number: A64520; MUID:97394467; PMID:9252185
 A;Accession: H64694
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-842 <TOM>
 A;Cross-references: UNIPROT:O25950; GB:AB000640; GB:AE000511; NID:92314572; PIDN:AAD0844
 C;Genetics:
 A;Start codon: GTG
 C;Superfamily: ferriochrome-iron receptor; tonB-dependent receptor amino-terminal homolog

 Query Match 30.4%; Score 49; DB 2; Length 842;
 Best Local Similarity 60.0%; Pred.No.:73;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

 QY 6 QPKPNNSSYAYNYFT 20
 |||||
 Db 827 QPAPGRSVTYLNYT 841

Search completed: November 10, 2004, 15:55:06
 Job time : 11.2491 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 50.2847 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-69

Perfect score: 151
Sequence: 1 EANKQPKPNNSTAYNFTGVSIPLSYKP 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	36.6	771	Q9LFA9	Q9LFA9 arabidopsis
2	55.5	34.5	351	Q934G7	Q934G7 bruceella su
3	55	34.2	423	Q82EL3	Q82EL3 streptomyc
4	54.5	33.9	183	Q6MM10	Q6MM10 bdellovibri
5	54.5	33.9	183	CAE79696	CAE79696 bdellovib
6	54	33.5	271	Q951G1	Q951G1 gadus morhu
7	54	33.5	609	Q923X7	Q923X7 pseudomonas
8	54	33.5	717	MYC_DROME	Q9W4S7 drosophila
9	53	32.9	304	Q89HH5	Q89HH5 bradyrhizob
10	53	32.9	321	Q8TA73	Q8TA73 eimeria nec
11	53	32.9	351	Q728T7	Q728T7 desulfovibr
12	53	32.9	351	Q8S6987	Q8S6987 desulfovi
13	53	32.9	411	Q86128	Q86128 dictyosteli
14	53	32.9	423	Q811B1	Q811B1 bacillus ce
15	53	32.9	1766	Q70H33	Q70H33 fowlpox vir
16	53	32.9	1766	Q9J599	Q9J599 fowlpox vir
17	53	32.9	1766	CAE52664	CAE52664 fowlpox v
18	52	32.3	158	Q6GBW7	Q6GBW7 staphylococ
19	52	32.3	158	Q6GJ66	Q6GJ66 staphylococ
20	52	32.3	158	Q8NXU4	Q8NXU4 staphylococ
21	52	32.3	158	Q99W10	Q99W10 staphylococ
22	52	32.3	158	Q7A745	Q7A745 staphylococ
23	52	32.3	228	Q9YEA4	Q9YEA4 aeropyrum p
24	52	32.3	230	Q96E51	Q96E51 homo sapien
25	52	32.3	506	Q8L341	Q8L341 vibrio chol
26	51.5	32.0	119	Q935Q3	Q935Q3 salmonella
27	51.5	32.0	3262	Q8YT08	Q8YT08 anabaena sp
28	51	31.7	230	RUM1_SCHPO	P40380 schizosacch
29	51	31.7	329	Q95158	Q95158 gadus morhu
30	51	31.7	864	AGLU_MUGJA	Q92442 mucor javan
31	51	31.7	1702	Q8DL21	Q8DL21 synechococc

Q8I2P8 plasmodium
Q8YD17 bruceella me
Q8T3Z6 drosophila
Q8FVA8 bruceella su
Q9VCG6 drosophila
Q8PN90 xanthomonas
Q98273 gadus morhu
P94168 actinobacil
Q7VH57 helicobacte
Q8GGJ9 fugu rubrip
Q6CVT9 kluyveromyc
Q9AQF3 clostridium
P32571 saccharomyc
O77322 plasmodium

ALIGNMENTS

RESULT 1

Q9LFA9 PRELIMINARY; PRT; 771 AA.
ID Q9LFA9
AC Q9LFA9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F7J8_250.
GN Name=F7J8_250;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,
RA Wohldmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI3189; CAB69855.1; -
DR PIR; T45967; T45967.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR004274; NIF.
DR Pfam; PF00035; dsrm; 1.
DR SMART; SM00577; CPDC; 1.
DR SMART; SM00358; DSRM; 1.
DR PROSITE; PS00137; DS_RBD; 1.
KW Hypothetical protein_
SQ SEQUENCE 771 AA; 86025 MW; 0C21587E2FF6491D CRC64;

Query Match 36.6%; Score 59; DB 2; Length 771;

Best Local Similarity 41.7%; Pred. No. 16;

Matches 15; Conservative 5; Mismatches 6; Indels 10; Gaps 2;

QY 1 EANKQPKPNNSTAYNFTGVSIPLSYKP 30

Db 477 EOKPTEPKQIAVIENNASTA---TAAALLPSHAP 508

RESULT 2

Q934G7 PRELIMINARY; PRT; 351 AA.
ID Q934G7
AC Q934G7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sugar-binding protein precursor.

GN Name=chvE;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330;
RX MEDLINE=21405735; PubMed=11514518;
RA Alvarez-Martinez M.T., Machold J., Weise C., Schmidt-Eisenlohr H.,
Baron C., Rouot B.;
RT "Brucella suis genome of the Agrobacterium tumefaciens chromosomal
virulence gene chvE is essential for sugar utilization but not for
survival in macrophages";
RT J. Bacteriol. 183:5343-5351(2001).
RL EMBL; AU305234; CAC67776.1; -;
DR InterPro; IPR001761; PeriplaBP/Lac1.
DR Pfam; PF00532; Peripla_BP_1; 1.
KW Signal.
FT SIGNAL
FT CHAIN
FT SIGNAL
SQ SEQUENCE 351 AA; 38077 MW; E5566482F1F3132A CRC64;
Query Match 34.5%; Score 55.5; DB 2; Length 351;
Best Local Similarity 48.3%; Pred. No. 20;
Matches 14; Conservative 6; Mismatches 4; Indels 5; Gaps 3;
QY 4 RKQPKNNSTAYNFTGVSLPSV--KP 30
DB 302 RKEPVNDTKT--YN-NGVKVPSYLLKP 327
RESULT 3
Q82EL3 PRELIMINARY; PRT; 423 AA.
AC Q82EL3
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SAV4601;
OS Streptomyces avermitilis;
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AF005039; BAC72313.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 423 AA; 46345 MW; CDC14CA678A3517D CRC64;
Query Match 34.2%; Score 55; DB 2; Length 423;
Best Local Similarity 43.8%; Pred. No. 30;
Matches 14; Conservative 4; Mismatches 8; Indels 6; Gaps 2;

QY 4 RKQPKNNSTAYNFTGVSLPSV-----YKP 30
DB 102 RKVAKANNASTYF--LSGYLLPEKRDLYKP 132
RESULT 4
Q6MM10 PRELIMINARY; PRT; 183 AA.
AC Q6MM10
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyltransferase (EC 2.3.1.-).
GN OrderedLocustNames=Bdi838;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
Sokkett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
genomic perspective";
RL Science 303:689-692(2004).
DR EMBL; BX842651; CAE79696.1; -;
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0018740; F:transferase activity; IEA.
DR InterPro; IPR000182; GCNSacetyltrans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Acyltransferase; Complete.proteome; Transferase.
SQ SEQUENCE 183 AA; 20648 MW; 9742371BF55E0973 CRC64;
Query Match 33.9%; Score 54.5; DB 2; Length 183;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 2; Mismatches 6; Indels 3; Gaps 1;
QY 9 PNNSSSTAY---YNFTGVSLPSY 28
DB 135 PNNSSIRYTKYGTSTVATHPNY 157
RESULT 5
CAE79696 PRELIMINARY; PRT; 183 AA.
AC CAE79696
DT 02-VAR-2004 (TrEMBLrel. 27, Created)
DT 02-VAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-VAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyltransferase (EC 2.3.1.-).
GN Bdi838.
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
Sokkett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
genomic perspective";
RL Science 303:689-692(2004).
DR EMBL; BX842651; CAE79696.1; -;
KW Acyltransferase; Transferase.
SQ SEQUENCE 183 AA; 20648 MW; 9742371BF55E0973 CRC64;
Query Match 33.9%; Score 54.5; DB 2; Length 183;

Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 9 PNNSSSTAY---YNTGVSILPSY 28

Db 135 PNNSSIRYLTKYGTSTVATHNY 157

RESULT 6

ID Q95161 PRELIMINARY; PRT; 271 AA.
AC Q95161
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MHC Class Ia antigen (Fragment).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]

SEQUENCE FROM N.A.
RP MEDLINE=21850511; PubMed=11862395;
RA Miller K.M., Kaukinen K.H., Schulze A.D.;
RT "Expansion and contraction of major histocompatibility complex genes:
a teleostean example.";

EL Immunogenetics 53:941-963(2002).
DR EMBL; AF414217; AAL14543.1; -.

DR GO; GO:0016020; C:Membrane; IEA.
DR GO; GO:0006955; P:Immune response; IEA.

DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG cl.

DR InterPro; IPR001039; MHC I.
DR Pfam; PF07654; CI-set; 1.

DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.

DR SMART; SM00407; IGCI; 1.
DR PROSITE; PS00835; IG_LIKE; 1.

FT NON TER 1
SQ SEQUENCE 271 AA; 31112 MW; EDB1E7C64AD01FBC CRC64;

Query Match 33.5%; Score 54; DB 2; Length 271;

Best Local Similarity 38.5%; Pred. No. 25;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 KQPKPNNSSSTAYVNTGVSILPSYKP 30

Db 240 KRRPDDSDTSSENTEGQKLAPEYQP 265

RESULT 7

ID Q923X7 PRELIMINARY; PRT; 609 AA.
AC Q923X7

DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Endoglucanase.

OS Pseudomonas sp. YD-15.
OC Bacteria; Proteobacteria.

OX NCBI_TaxID=72609;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=YD-15;

RX MEDLINE=20128368; PubMed=10664983;
RA Her S., Lee H.-S., Choi S.-J., Choi S.-W., Choi H.-J., Yoon S.-S.,

RA "Cloning and sequencing of beta-1,4-endoglucanase gene (celA) from
RT Pseudomonas sp. YD-15.";

RL Lett. Appl. Microbiol. 29:389-395(1999).
DR EMBL; AF033262; AAD01959.1; -.

DR HSSP; P04954; 1CLC.

DR GO; GO:0008810; F:cellulase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001701; Glyco_hydro_9.
DR InterPro; IPR004197; Glyco_hydro_9ig.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR InterPro; IPR000566; Lipocin_cytFABP.
DR Pfam; PF02927; Celd_N; 1.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
DR PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN 1.
SQ SEQUENCE 609 AA; 63618 MW; 7885B357203351A3 CRC64;

Query Match 33.5%; Score 54; DB 2; Length 609;

Best Local Similarity 56.2%; Pred. No. 63;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ANRKQPKPNNSSSTAY 17

Db 451 ANQAYPAPNNACTVY 466

RESULT 8

MYC DROME

ID MYC DROME STANDARD; PRT; 717 AA.
AC Q9W457; O96903; P91665;

DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Myc protein (dMyc1) (diminutive protein).

GN Name=dm; ORFNames=CG10798;
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.; FUNCTION, SUBUNIT, TISSUE SPECIFICITY, AND
RP DEVELOPMENTAL STAGE.

RC STRAIN=Oregon-R;
RX MEDLINE=97078158; PubMed=8929412;

RA Gallant P., Shiao Y., Cheng P.F., Parkhurst S.M., Eisenman R.N.;
RT "Myc and Max homologs in Drosophila.";

RL Science 274:1523-1527(1996).
RN [2]

RP REVISION TO 274.
RA Gallant P., Shiao Y., Cheng P.F., Parkhurst S.M., Eisenman R.N.;

RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX PubMed=9037036;

RA Schreiber-Agus N., Stein D., Chen K., Goltz J.S., Stevens L.,
RA Depinho R.A.;

RT "Drosophila Myc is oncogenic in mammalian cells and plays a role in
RT the diminutive phenotype.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:1235-1240(1997).
RN [4]

RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Wan K.H., Doyle C., Baxter E.G., Blazek R.G., Champe M., Pfeiffer B.D.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,


```

RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BAC51281.1; -.
DR InterPro; IPR005628; GspK.
DR Pfam; PF03934; GspK; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 32633 MW; D576789684D7D803 CRC64;

Query Match 32.9%; Score 53; DB 2; Length 304;
Best Local Similarity 40.0%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy 6 QPKPNNSTAYNFTGVSIPLSYKP 30
Db 144 EPGQDNPEADYRTLGASLYLPRHAP 168

RESULT 10
Q8TA73 PRELIMINARY; PRT; 321 AA.
AC Q8TA73;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Npm219 protein.
GN Name=Npm219;
OS Eimeria necatrix.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Eimeria.
OX NCBI_TaxID=51315;
RN [1]_TaxID=51315;
RP SEQUENCE FROM N.A.
RX MEDLINE=22768932; PubMed=12887190;
RA Tajima O., Onaga H., Nakamura T.;
RT "An enzyme-linked immunosorbent assay with the recombinant merozoite
RT protein as antigen for detection of antibodies to Eimeria necatrix.";
RL Avian Dis. 47:309-318(2003).
DR EMBL; AB070239; BAB85126.1; -.
DR EMBL; AB070239; BAB85126.1; -.
SQ SEQUENCE 321 AA; 34175 MW; 10CB1CA4E32C1999 CRC64;

Query Match 32.9%; Score 53; DB 2; Length 321;
Best Local Similarity 45.0%; Pred. No. 42;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ANRKQPKPNNSTAYNFTG 21
Db 109 ANRRKKPDPTPAAAIYDFVG 128

RESULT 11
Q728T7 PRELIMINARY; PRT; 351 AA.
AC Q728T7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HD domain protein.
GN OrderedLocusNames=DVU2515;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]_TaxID=882;
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Koonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,

```

```

RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Kouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AB017317; AAS96987.1; -.
DR TIGR; DVU2515; -.
DR InterPro; IPR006574; HD_hydro.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Complete proteome.
SQ SEQUENCE 351 AA; 39994 MW; EB3DE1E08A05E58C CRC64;

Query Match 32.9%; Score 53; DB 2; Length 351;
Best Local Similarity 39.3%; Pred. No. 47;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

Qy 5 KOPKNNSTAYNFTG--VSILPSYKP 30
Db 5 KSPIDNISEEYQISAILSSFPKYP 32

RESULT 12
AAS96987 PRELIMINARY; PRT; 351 AA.
ID AAS96987;
AC AAS96987;
DT 27-APR-2004 (TrEMBLrel. 27, Created)
DT 27-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE HD domain protein.
GN DVU2515;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]_TaxID=882;
RP SEQUENCE FROM N.A.
RX PubMed=15077118;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Kouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AB017317; AAS96987.1; -.
DR TIGR; DVU2515; -.
DR InterPro; IPR006574; HD_hydro.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Complete proteome.
SQ SEQUENCE 351 AA; 39994 MW; EB3DE1E08A05E58C CRC64;

Query Match 32.9%; Score 53; DB 2; Length 351;
Best Local Similarity 39.3%; Pred. No. 47;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

Qy 5 KOPKNNSTAYNFTG--VSILPSYKP 30
Db 5 KSPIDNISEEYQISAILSSFPKYP 32

RESULT 13
Q86128 PRELIMINARY; PRT; 411 AA.
ID Q86128;
AC Q86128;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)

```

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE VARV Bangladesh B22R orthologue.
 GN Name=fp9.123;
 OS Fowlpox virus (isolate HP-438[Munich]).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Avipoxvirus.
 OC NCBI_TaxID=10263;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Skinner M.A., Laidlaw S.M.;
 RT "Comparison of the genome sequence of FP9, an attenuated, tissue
 RT culture-adapted European fowlpox virus, with those of virulent
 RT American and European viruses";
 RL J. Gen. Virol. 85:305-322(2004).
 DR EMBL; AJ581527; CAE52664.1;
 DR InterPro; IPR000209; Pept_S8_S53.
 DR InterPro; IPR007490; Poxvirus_B22R.
 DR Pfam; PF04395; Poxvirus_B22R; 1.
 DR PROSITE; PS00136; SUBTILASE ASP; UNKNOWN 1.
 SQ SEQUENCE 1766 AA; 201145 MW; B3D87FF4A52679D9 CRC64;

Query Match 32.9%; Score 53; DB 2; Length 1766;
 Best Local Similarity 47.8%; Pred. No. 2.9e+02;
 Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 8 KPNNSSTAYNFTGVSIILPSYKP 30
 Db 205 KSKNKSTITVTENGVSPPYEP 227

Search completed: November 10, 2004, 15:53:32
 Job time : 52.3616 secs

DE Similar to Dictyostelium discoideum (Slime mold). MkpA protein.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OC NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910;
 RA Gloeckner G., Eichinger L., Safranski K., Pachebat J., Dear P.,
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tuncgall B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RL Nature 418:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Baumgart C.;
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC116305; AAO52322.1; -
 SQ SEQUENCE 411 AA; 46427 NW; A0974D03F43FB6F1 CRC64;

Query Match 32.9%; Score 53; DB 2; Length 411;
 Best Local Similarity 43.5%; Pred. NO. 56;
 Matches 10; Conservative 5; Mismatches 2; Indels 6; Gaps 1;

QY 14 TAYNFTGVSI-----LPSYKP 30
 Db 224 TAFNFTNSNVKFSDCSIPSYKP 246

RESULT 14
 Q8LIB1
 ID Q8LIB1 PRELIMINARY; PRT; 423 AA.
 AC Q8LIB1;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Penicillin-binding protein.
 GN ORFNames=BC0483;
 OS Bacillus cereus (strain ATCC 14579 / DSM 31).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=226900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
 RX Ivanova N., Sorokin A., Anderson I., Galleron N., Candellon B.,
 RA Kapatal V., Bhattacharyya A., Reznik N., Mikhailova N., Lapidus A.,
 RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
 RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
 RA Overbeek R., Kyrpides N.C.;
 RT "Genome sequence of Bacillus cereus and comparative analysis with
 RT Bacillus anthracis.";
 RL Nature 423:87-91(2003).
 DR EMBL; AE016999; AAP07521.1; -
 DR InterPro; IPR001466; Beta_lactamase.
 DR Pfam; PF00144; Beta_lactamase; 1.
 SQ SEQUENCE 423 AA; 47619 MW; 19ABF50A949B37E9 CRC64;

Query Match 32.9%; Score 53; DB 2; Length 423;
 Best Local Similarity 56.2%; Pred. NO. 58;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANKQPKNNSTAY 17
 Db 145 ANKQOVNNSETTF 160

RESULT 15
 Q7OH33
 ID Q7OH33 PRELIMINARY; PRT; 1766 AA.
 AC Q7OH33;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 7.40214 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-70

Perfect score: 85
Sequence: 1 GSLTHNNIKPSSTR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/6C COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	89.4	459	3	US-09-097-889-22
2	76	89.4	459	4	US-09-098-079-22
3	47	55.3	163	4	US-09-248-796A-27380
4	42	49.4	108	3	US-09-187-859-6
5	42	49.4	108	4	US-09-839-542B-6
6	42	49.4	108	4	US-09-535-852-6
7	42	49.4	780	1	US-08-188-228-50
8	42	49.4	780	1	US-08-332-843-44
9	42	49.4	780	1	US-08-332-638-50
10	41	48.2	271	4	US-09-248-796A-14668
11	41	48.2	510	3	US-08-246-489-2
12	40	47.1	50	4	US-09-621-976-7482
13	40	47.1	128	4	US-09-489-039A-14108
14	40	47.1	128	4	US-09-513-999C-5518
15	40	47.1	165	4	US-09-270-767-36974
16	40	47.1	165	4	US-09-270-767-52191
17	40	47.1	600	2	US-08-821-119-19
18	40	47.1	600	2	US-08-821-118-2
19	39	45.9	148	4	US-09-248-796A-21587
20	39	45.9	179	4	US-09-270-767-31777
21	39	45.9	179	4	US-09-270-767-46994
22	39	45.9	244	4	US-09-270-767-46968
23	39	45.9	1461	4	US-09-976-594-531
24	38	44.7	138	4	US-09-489-039A-7633
25	38	44.7	159	4	US-09-508-691-5
26	38	44.7	225	4	US-09-540-236-2212
27	38	44.7	261	4	US-09-602-787A-478

28 38 44.7 434 4 US-09-489-039A-11674 Sequence 11674, A
29 38 44.7 464 4 US-09-711-164-430 Sequence 430, App
30 38 44.7 526 4 US-09-248-796A-18475 Sequence 18475, A
31 38 44.7 570 3 US-09-134-001C-2972 Sequence 2972, Ap
32 38 44.7 572 4 US-09-248-796A-19066 Sequence 19066, A
33 38 44.7 660 4 US-09-248-796A-19966 Sequence 19966, A
34 38 44.7 811 4 US-09-248-796A-19965 Sequence 19965, A
35 38 44.7 832 4 US-09-758-282B-251 Sequence 251, App
36 38 44.7 832 4 US-09-758-282B-268 Sequence 268, App
37 38 44.7 832 4 US-09-577-304A-251 Sequence 251, App
38 38 44.7 832 4 US-09-577-304A-268 Sequence 268, App
39 38 44.7 838 4 US-09-758-282B-261 Sequence 261, App
40 38 44.7 838 4 US-09-758-282B-265 Sequence 265, App
41 38 44.7 838 4 US-09-577-304A-261 Sequence 261, App
42 38 44.7 838 4 US-09-577-304A-265 Sequence 265, App
43 38 44.7 1154 4 US-09-134-000C-6122 Sequence 6122, Ap
44 38 44.7 1321 2 US-08-317-310A-64 Sequence 64, Appl
45 37 43.5 17 4 US-09-424-840-64 Sequence 64, Appl

ALIGNMENTS

RESULT 1

US-09-097-889-22

; Sequence 22, Application US/09097889

; Patent No. 6218117

; GENERAL INFORMATION:

; APPLICANT: HerinStadt, Corrina

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Davis, Robert E.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING

; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED AND BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/097,889

; FILING DATE: 15-JUN-1998

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Roseman Ph.D., Stephen J.

; REGISTRATION NUMBER: 43,058

; REFERENCE/DOCKET NUMBER: 660088.417

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 459 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-09-097-889-22

Query Match 89.4%; Score 76; DB 3; Length 459;

Best Local Similarity 87.5%; Pred. No. 0.00019;

Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHNNIKPSSTR 16

DB 417 GSLTHNNIKPSSTR 432

```

; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27380

Query Match 55.3%; Score 47; DB 4; Length 163;
Best Local Similarity 69.2%; Pred. No. 3,4;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSS 14
   |||||
Db 36 SYTHSNHKKPS 48

RESULT 4
US-09-187-859-6
; Sequence 6, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-187-859-6

Query Match 49.4%; Score 42; DB 3; Length 108;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSSR 16
   |||||
Db 20 SLPHVGVKIKSSVSR 34

RESULT 5
US-09-839-542B-6
; Sequence 6, Application US/09839542B
; Patent No. 6563996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-542B-6

Query Match 49.4%; Score 42; DB 4; Length 108;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSSR 16
   |||||
Db 20 SLPHVGVKIKSSVSR 34

RESULT 6

```



```

US-09-535-852-6
; Sequence 6, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-535-852-6

Query Match          49.4%; Score 42; DB 4; Length 108;
Best Local Similarity 53.3%; Pred. No. 14;
Matches      8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 SLTHHNNIKPSSTR 16
      |||||
Db      20 SLPHVGVKTKSSVSR 14

```

Qy 2 SLTHHNNIKPSSTR 16
|||: |||: |
Db 20 SLPHFVGKIKSSYSR 34

RESULT 7
US-08-188-228-50
; Sequence 50, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 APR 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 559772sand, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 780 amino acids
; TYPE: aminc acid
; TOPOLOGY: linear

TELEPHONE: (312) 474-8300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 780 amino acids
TYPE: amino acid
TOPOLOGY: linear

```

; MOLECULE TYPE: protein
US-08-188-228-50

Query Match 49.4%; Score 42; DB 1; Length 780;
Best Local Similarity 53.3%; Pred No. 1.2e-02;
Matches 8; Conservative 2; Mismatches 5; Indels

QY 2 SLTHHNINKPSSTR 16
   |||: |||:|
Db 59 SLPHVGVKIKSVSR 73

RESULT 8
US-08-332-643-44
; Sequence 44, Application US/08332643
; Patent No. 5639634
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-132-643-44

```

```

Query Match      49.4%; Score 42; DB 1; Length 780;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY      2 SLTHHNNIKPSSTR 16
Db      59 SLPHVGVKISVSR 73

```

RESULT 9
US-08-332-638-50
; Sequence 50, Application US/08332638
; Patent No. 5646250
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Borun
 ; CITY: Chicago
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/332,638
 ; FILING DATE: 01-NOV-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/872,643
 ; FILING DATE: 17 APR 1992
 ; APPLICATION NUMBER: US/08/049,460
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5646250and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 31340
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 50:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 780 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-332-638-50

Query Match 49.4%; Score 42; DB 1; Length 780;
 Best Local Similarity 53.3%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSSTR 16
 ||| : ||| :
 Db 59 SLPHVGVKIKSVSR 73

RESULT 10
 US-09-248-796A-14668
 ; Sequence 14668, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 14668
 ; LENGTH: 271
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-14668
 Query Match 48.2%; Score 41; DB 4; Length 271;
 Best Local Similarity 53.3%; Pred. No. 56;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GSLTHHNNIKPSST 15
 ||| : ||| :
 Db 153 GSITHTVIENKPSS 167
 RESULT 11
 US-08-246-489-2
 ; Sequence 2, Application US/08246489
 ; Patent No. 6225049
 ; GENERAL INFORMATION:
 ; APPLICANT: Lan, Michael S.
 ; APPLICANT: No. 6225049Kins, Abner L.
 ; TITLE OF INVENTION: NOVEL HUMAN INSULINOMA-ASSOCIATED CDNA
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson & Bear
 ; STREET: 620 Newport Center Drive
 ; CITY: Newport Beach
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/246,489
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/901,715
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned A.
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER: NIH012.012A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 510 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; US-08-246-489-2

Query Match 48.2%; Score 41; DB 3; Length 510;
 Best Local Similarity 57.1%; Pred. No. 1.1e+02;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 LTHHNNIKPSSTR 16
 ||| : ||| :
 Db 484 LTRHINKCHPSNR 497

RESULT 12
 US-09-621-976-7482
 ; Sequence 7482, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 38.6619 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-70

Perfect score: 85

Sequence: 1 GSLTHINNKPSTR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA: *

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/FCI_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/FCIUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	16	14	US-10-092-750-70
2	79	92.9	114	17	US-10-425-115-336967
3	76	89.4	459	9	US-09-098-079-22
4	76	89.4	459	15	US-10-428-487-36
5	76	89.4	459	16	US-10-408-765A-1
6	76	89.4	459	16	US-10-408-765A-2682
7	66	77.6	83	15	US-10-424-599-278981
8	49	57.6	123	16	US-10-437-963-111254
9	47	55.3	312	15	US-10-424-599-249656
10	45	52.9	235	15	US-10-282-122A-53894
11	44	51.8	42	17	US-10-425-115-259437
12	44	51.8	310	16	US-10-437-963-188297
13	44	51.8	314	15	US-10-424-599-225540

14	51.8	348	15	US-10-424-599-256084	Sequence 256084,
15	51.8	383	15	US-10-424-599-184641	Sequence 184641,
16	51.8	386	15	US-10-425-114-52425	Sequence 52425, A
17	51.8	394	15	US-10-425-114-36970	Sequence 36970, A
18	51.8	840	17	US-10-425-115-213418	Sequence 213418,
19	51.8	847	15	US-10-425-114-36827	Sequence 36827, A
20	51.8	1096	16	US-10-408-765A-747	Sequence 747, App
21	50.6	43	9	US-09-925-299-1211	Sequence 1211, App
22	50.6	43	10	US-09-925-299-1211	Sequence 1211, App
23	50.6	109	17	US-10-424-599-195180	Sequence 195180,
24	50.6	109	17	US-10-739-930-9987	Sequence 9987, App
25	50.6	291	14	US-10-369-493-19454	Sequence 19454, A
26	50.6	307	15	US-10-424-599-239832	Sequence 239832,
27	50.6	468	15	US-10-424-599-239830	Sequence 239830,
28	50.6	636	15	US-10-108-260A-2860	Sequence 2860, App
29	50.6	722	14	US-10-094-749-2220	Sequence 2220, App
30	49.4	47	17	US-10-425-115-270957	Sequence 270957,
31	49.4	95	15	US-10-424-599-251720	Sequence 251720,
32	49.4	108	14	US-10-006-869-6	Sequence 6, Appli
33	49.4	108	14	US-10-395-032-6	Sequence 6, Appli
34	49.4	108	16	US-10-714-564A-1372	Sequence 1372, Ap
35	49.4	374	15	US-10-424-599-242398	Sequence 242398,
36	49.4	493	15	US-10-424-599-256083	Sequence 256083,
37	49.4	566	14	US-10-029-386-32217	Sequence 32217, A
38	49.4	566	15	US-10-108-260A-3846	Sequence 3846, Ap
39	49.4	780	16	US-10-778-146-8	Sequence 8, Appli
40	49.4	784	9	US-09-905-983-52	Sequence 52, Appl
41	49.4	784	14	US-10-021-660-99	Sequence 99, Appl
42	49.4	784	14	US-10-174-677-23	Sequence 23, Appl
43	49.4	784	15	US-10-211-462-105	Sequence 105, App
44	49.4	897	14	US-10-138-927-45	Sequence 45, Appl
45	49.4	897	14	US-10-430-011-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-10-092-750-70
; Sequence 70, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XI
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-70

Query Match 100.0%; Score 85; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLTHINNKPSTR 16

Db 1 GSLTHINNKPSTR 16

RESULT 2

US-10-425-115-336967
; Sequence 336967, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(53222)B
 CURRENT APPLICATION NUMBER: US/10/425,115
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 369326
 SEQ ID NO 336967
 LENGTH: 114
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(114)
 OTHER INFORMATION: unsure at all Xaa locations
 FEATURE:
 OTHER INFORMATION: Clone ID: MRT4577_70494C.1.pep
 US-10-425-115-336967

Query Match 92.9%; Score 79; DB 17; Length 114;
 Best Local Similarity 93.8%; Pred. No. 6.4e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
 |||||
 DB 91 GSLTHHNNKPSSTR 106
 |||||

RESULT 3
 US-09-098-079-22
 Sequence 22, Application US/09098079
 Patent No. US20020064773A1
 GENERAL INFORMATION:
 APPLICANT: Herznstadt, Corrina
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Clevenger, William
 APPLICANT: Fahy, Eoin P.
 APPLICANT: Davis, Robert E.
 TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
 TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED AND BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/098,079
 FILING DATE: 15-JUN-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Rosenman Ph.D., Stephen J.
 REGISTRATION NUMBER: 43,058
 REFERENCE/DOCKET NUMBER: 660088.416
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear

US-09-098-079-22
 Query Match 89.4%; Score 76; DB 9; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00086;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GSLTHHNNKPSSTR 16
 |||||
 DB 417 GSLTHHNNKPSSTR 432
 |||||
 RESULT 4
 US-10-428-487-36
 Sequence 36, Application US/10428487
 Publication No. US20040006780A1
 GENERAL INFORMATION:
 APPLICANT: RASTELLI, LUCA K.
 APPLICANT: GERBER, HANS-PETER
 TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
 FILE REFERENCE: 09800080-0103
 CURRENT APPLICATION NUMBER: US/10/428,487
 CURRENT FILING DATE: 2003-05-02
 PRIOR APPLICATION NUMBER: 09/815,153
 PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,201
 PRIOR FILING DATE: 2000-03-22
 NUMBER OF SEQ ID NOS: 84
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 36
 LENGTH: 459
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-428-487-36

Query Match 89.4%; Score 76; DB 15; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00086;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
 |||||
 DB 417 GSLTHHNNKPSSTR 432
 |||||

RESULT 5
 US-10-408-765A-1
 Sequence 1, Application US/10408765A
 Publication No. US20040101874A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Eoin D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary M.
 APPLICANT: Warnock, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 FILE REFERENCE: 660088.465
 CURRENT APPLICATION NUMBER: US/10/408,765A
 CURRENT FILING DATE: 2003-04-04
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 459
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-408-765A-1

Query Match 89.4%; Score 76; DB 16; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00086;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16

Db 417 GSLTHHNNKPSFTR 432
|||||:|:|:|

RESULT 6

US-10-408-765A-2682
; Sequence 2682, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2682
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2682

Query Match 89.4%; Score 76; DB 16; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00086;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSLTHHNNKPSSTR 16
|||||:|:|:|
Db 417 GSLTHHNNKPSFTR 432

RESULT 7

US-10-424-599-278981
; Sequence 278981, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 278981
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(83)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_93943C.1.pep
US-10-424-599-278981

Query Match 77.6%; Score 66; DB 15; Length 83;
Best Local Similarity 75.0%; Pred. No. 0.0057;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSLTHHNNKPSSTR 16
|||||:|:|:|
Db 59 GSVSHHNNKPSFTR 74

RESULT 8

US-10-437-963-111254
; Sequence 111254, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111254
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(123)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1524C.1.pep
US-10-437-963-111254

Query Match 57.6%; Score 49; DB 16; Length 123;
Best Local Similarity 60.0%; Pred. No. 4.8;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GSLTHHNNKPSST 15
|||||:|:|:|
Db 5 GSLTYRIGSVSPST 19

RESULT 9

US-10-424-599-249656
; Sequence 249656, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249656
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67469C.1.pep
US-10-424-599-249656

Query Match 55.3%; Score 47; DB 15; Length 312;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLTHHNNKPSST 15
|||||:|:|:|
Db 34 SLSHRLNAVKPSKT 47

```
RESULT 10
US-10-282-122A-53894
; Sequence 53894, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chisen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 53894
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53894

Query Match 52.9%; Score 45; DB 15; Length 235;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GSLTHHNNIKPSSSTR 16
Db 190 GALLDEINRIKPSAK 205

RESULT 11
US-10-425-115-258437
; Sequence 258437, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28

US-10-425-115-258437
; Sequence 53894, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chisen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 53894
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-53894

Query Match 52.9%; Score 45; DB 15; Length 235;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GSLTHHNNIKPSSSTR 16
Db 190 GALLDEINRIKPSAK 205

RESULT 11
US-10-425-115-258437
; Sequence 258437, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258437
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167281C.1.pep
US-10-425-115-258437

Query Match 51.8%; Score 44; DB 17; Length 42;
Best Local Similarity 60.0%; Pred. No. 9.6;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSSSTR 16
Db 1 SHTHFQIRPSSR 15

RESULT 12
US-10-437-963-188297
; Sequence 188297, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188297
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84915C.1.pep
US-10-437-963-188297

Query Match 51.8%; Score 44; DB 16; Length 310;
Best Local Similarity 61.5%; Pred. No. 82;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLTHHNNIKPSS 14
Db 28 SLQHRNNITPAA 40

RESULT 13
US-10-424-599-225540
; Sequence 225540, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 225540
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Glycine max
```


FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_45694C.1.pep
US-10-424-599-225540

Query Match 51.8%; Score 44; DB 15; Length 314;
Best Local Similarity 72.7%; Pred. No. 83;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLTHHNNIKP 12
Db 114 STTHNNISP 124

RESULT 14
US-10-424-599-256084
Sequence 256084, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 256084
LENGTH: 348
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_73265C.1.pep
US-10-424-599-256084

Query Match 51.8%; Score 44; DB 15; Length 348;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LTHHNNIKPSS 14
Db 262 LTHHYNLGPSS 273

RESULT 15
US-10-424-599-184641
Sequence 184641, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 184641
LENGTH: 383
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_137747C.1.pep
US-10-424-599-184641

Query Match 51.8%; Score 44; DB 15; Length 383;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LTHHNNIKPSS 14

Db 53 LTHHLNRLKSSA 64

Search completed: November 11, 2004, 07:41:46
Job time : 38.6619 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 10, 2004, 14:52:32 ; Search time 5.46619 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-70
Perfect score: 85
Sequence: 1 GSLTHINNKPSTR 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	89.4	152	2 A00435	NADH2 dehydrogenas
2	76	89.4	459	1 DNHUN4	NADH2 dehydrogenas
3	73	85.9	459	2 T17141	NADH2 dehydrogenas
4	67	78.8	459	1 QXGI4M	NADH2 dehydrogenas
5	66	77.6	152	2 A00437	NADH2 dehydrogenas
6	65	76.5	152	2 A00436	NADH2 dehydrogenas
7	65	76.5	459	2 A59134	NADH2 dehydrogenas
8	65	76.5	459	2 T11372	NADH2 dehydrogenas
9	65	76.5	459	2 T11866	NADH2 dehydrogenas
10	64	75.3	459	2 T11256	NADH2 dehydrogenas
11	62	72.9	459	2 B58851	NADH2 dehydrogenas
12	62	72.9	459	2 T10981	NADH2 dehydrogenas
13	61	71.8	459	2 T11029	NADH2 dehydrogenas
14	60	70.6	459	2 T11502	NADH2 dehydrogenas
15	60	70.6	459	2 S41829	NADH2 dehydrogenas
16	59	69.4	459	2 QXBO4M	NADH2 dehydrogenas
17	59	69.4	459	2 T11515	NADH2 dehydrogenas
18	59	69.4	459	2 T11411	NADH2 dehydrogenas
19	58	68.2	459	2 S41844	NADH2 dehydrogenas
20	58	68.2	459	2 S26160	NADH2 dehydrogenas
21	58	68.2	459	2 S04756	NADH2 dehydrogenas
22	56	65.9	152	2 I77325	NADH2 dehydrogenas
23	56	65.9	152	2 I77323	NADH2 dehydrogenas
24	56	65.9	152	2 I77321	NADH2 dehydrogenas
25	56	65.9	152	2 I77327	NADH2 dehydrogenas
26	56	65.9	459	2 T17137	NADH2 dehydrogenas
27	56	65.9	459	2 T17148	NADH2 dehydrogenas
28	56	65.9	459	2 T17152	NADH2 dehydrogenas
29	56	65.9	459	2 T17155	NADH2 dehydrogenas

30 56 65.9 459 2 T17358 NADH2 dehydrogenas
31 56 65.9 459 2 T17364 NADH2 dehydrogenas
32 55 64.7 459 2 T11149 NADH2 dehydrogenas
33 55 64.7 459 2 T17349 NADH2 dehydrogenas
34 55 64.7 459 2 T17184 NADH2 dehydrogenas
35 55 64.7 459 2 T17163 NADH2 dehydrogenas
36 55 64.7 459 2 T17172 NADH2 dehydrogenas
37 55 64.7 459 2 T17144 NADH2 dehydrogenas
38 55 64.7 459 2 T17175 NADH2 dehydrogenas
39 55 64.7 459 2 T17178 NADH2 dehydrogenas
40 55 64.7 459 2 T17181 NADH2 dehydrogenas
41 54 63.5 459 2 T11398 NADH2 dehydrogenas
42 53 62.4 152 2 I77318 NADH2 dehydrogenas
43 53 62.4 459 2 T17361 NADH2 dehydrogenas
44 52 61.2 459 2 T11489 NADH2 dehydrogenas
45 51 60.0 152 2 I57440 NADH2 dehydrogenas

ALIGNMENTS

RESULT 1
A00435
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - chimpanzee mitochondrion (fragm
N;Alternate names: NADH-ubiquinone oxidoreductase chain 4
C;Species: Mitochondrion Pan troglodytes (Chimpanzee)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A00435
R;Brown, W.M.; Prager, E.M.; Wang, A.; Wilson, A.C.
J. Mol. Evol. 18, 225-239, 1982
A;Title: Mitochondrial DNA sequences of primates: tempo and mode of evolution.
A;Reference number: A00435; MUID:82242101; PMID:6284948
A;Molecule type: DNA
A;Residues: 1-152 <BROW>
A;Cross-references: UNIPROT:P03906; GB:V00672; NID:g13369; PIDN:CAB51801.1; PID:g567962
C;Genetics: mitochondrion
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 89.4%; Score 76; DB 2; Length 152;
Best Local Similarity 87.5%; Pred No. 1.5e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSLTHINNKPSTR 16
||| ||||| |||||
Db 110 GSLTHINNKPSTR 125
RESULT 2
DNHUN4
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - human mitochondrion
N;Alternate names: NADH-ubiquinone oxidoreductase chain 4
C;Species: Mitochondrion Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004
C;Accession: A00434; B00435
R;Anderson, S.; Bankier, A.T.; Barrell, B.G.; de Bruijn, M.H.L.; Coulson, A.R.; Drouin,
Nature 290, 457-465, 1981
A;Title: Sequence and organization of the human mitochondrial genome.
A;Reference number: A00151; MUID:81173052; PMID:7219534
A;Accession: A00434
A;Molecule type: DNA
A;Residues: 1-459 <AND>
A;Cross-references: UNIPROT:P03905; GB:J01415; GB:M12548; GB:M58503; GB:M63932; GB:M639
R;Brown, W.M.; Prager, E.M.; Wang, A.; Wilson, A.C.
J. Mol. Evol. 18, 225-239, 1982
A;Title: Mitochondrial DNA sequences of primates: tempo and mode of evolution.
A;Reference number: A00435; MUID:82242101; PMID:6284948
A;Accession: B00435
A;Molecule type: DNA

```

A;Residues: 308-459 <BRO>
A;Cross-references: GB:L00016; EMBL:V00659; NID:G337302
A;Note: this ORF is not annotated in GenBank entry HUMTTRPR, release 106
C;Genetics:
A;Gene: GDB:MTND4
A;Cross-references: GDB:118914; OMIM:516003
A;Map position: MTH10760-12137
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match      89.4%; Score 76; DB 1; Length 459;
Best Local Similarity 87.5%; Pred. No. 5.1e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  GSLTHHNNKPSSTR 16
Db      417  GSLTHHNNKPSFTR 432

RESULT 3
T17141
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Colobus guereza mit
C;Species: mitochondrion Colobus guereza (guereza)
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: T17141
R;Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi
Int. J. Primatol. 18, 305-320, 1997
A;Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A;Reference number: Z18709
A;Accession: T17141
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <WAN>
A;Cross-references: UNIPROT:O21697; EMBL:U92950; NID:G2290413; PID:G2290416; PIDN:AAD046
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match      85.9%; Score 73; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00017;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GSLTHHNNKPSSTR 16
Db      417  GSLTHHNNKPSFTR 432

RESULT 4
OQGI4M
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - common gibbon mitochondrion (fra
C;Species: mitochondrion Hylobates lar (common gibbon, white-handed gibbon)
C;Date: 17-Dec-1982 #sequence_revision 26-Aug-1999 #text_change 09-Jul-2004
C;Accession: T11843; A00438
R;Aranson, U.; Gullberg, A.; Xu, X.
Hereditas 124, 185-189, 1996
A;Title: A complete mitochondrial DNA molecule of the white-handed gibbon, Hylobates lar
A;Reference number: Z17353
A;Accession: T11842
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-459 <ARN>
A;Cross-references: UNIPROT:P03909; EMBL:X99256; PIDN:CAA67637.1
A;Experimental source: isolate Ester
R;Brown, W.M.; Prager, E.M.; Wang, A.; Wilson, A.C.
J. Mol. Evol. 18, 225-239, 1982
A;Title: Mitochondrial DNA sequences of primates: tempo and mode of evolution.
A;Reference number: A00435; MUID:82242101; PMID:6284948
A;Accession: A00438
A;Molecule type: DNA
A;Residues: 308-459 <BRO>
A;Cross-references: GB:V00659; NID:G12996; PIDN:CAB51360.1; PID:G5579007
C;Genetics:

```

```

A;Genome: mitochondrion
A;Genetic code: SGCI
A;Note: NADH4
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match      79.8%; Score 67; DB 1; Length 459;
Best Local Similarity 75.0%; Pred. No. 0.0018;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1  GSLTHHNNKPSSTR 16
Db      417  GLTTHHNNKPSLTR 432

RESULT 5
A00437
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - orangutan mitochondrion (fragme
N;Alternate names: NADH-ubiquinone oxidoreductase chain 4
C;Species: mitochondrion Pongo pygmaeus (orangutan)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A00437
R;Brown, W.M.; Prager, E.M.; Wang, A.; Wilson, A.C.
J. Mol. Evol. 18, 225-239, 1982
A;Title: Mitochondrial DNA sequences of primates: tempo and mode of evolution.
A;Reference number: A00435; MUID:82242101; PMID:6284948
A;Accession: A00437
A;Molecule type: DNA
A;Residues: 1-152 <BRO>
A;Cross-references: UNIPROT:P03908; GB:V00675; NID:G13447; PIDN:CAB51802.1; PID:G567962
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match      77.6%; Score 66; DB 2; Length 152;
Best Local Similarity 75.0%; Pred. No. 0.00077;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1  GSLTHHNNKPSSTR 16
Db      110  GTPTHHNNKPSFTR 125

RESULT 6
A00436
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - gorilla mitochondrion (fragment
N;Alternate names: NADH-ubiquinone oxidoreductase chain 4
C;Species: mitochondrion Gorilla gorilla (gorilla)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A00436
R;Brown, W.M.; Prager, E.M.; Wang, A.; Wilson, A.C.
J. Mol. Evol. 18, 225-239, 1982
A;Title: Mitochondrial DNA sequences of primates: tempo and mode of evolution.
A;Reference number: A00435; MUID:82242101; PMID:6284948
A;Accession: A00436
A;Molecule type: DNA
A;Residues: 1-152 <BRO>
A;Cross-references: UNIPROT:P03907; GB:L00015; NID:G337102; PIDN:AAA69717.1; PID:G98626
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match      76.5%; Score 65; DB 2; Length 152;
Best Local Similarity 75.0%; Pred. No. 0.0011;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  GSLTHHNNKPSSTR 16
Db      110  GLTTHHNNKPSFTR 125

```

```
RESULT 7
A59154
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - western lowland gorilla mitochondrion
N/Alternate names: NADH-ubiquinone oxidoreductase chain 4
C/Species: Mitochondrion Gorilla gorilla gorilla (western lowland gorilla)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: A59154
R/Xu, X.; Arnason, U.
Mol. Biol. Evol. 13, 691-698, 1996
A/Title: A complete sequence of the mitochondrial genome of the Western lowland gorilla.
A/Reference number: Z17269; MUID:196212991; PMID:8676744
A/Accession: A59154
A/Status: preliminary; nucleic acid sequence not shown; translation not shown; translated
A/Molecule type: DNA
A/Residues: 1-459 <XUX>
A/Cross-references: UNIPROT:Q97974; GB:X93347; NID:gl304307; GSPDB:GN00106
A/Note: submitted to GenBank, November 1995
A/Note: this translation is not annotated in GenBank entry GGMITG, release 114.0
A/Note: the termination resulting from transcript polyadenylation is shown
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 76.5%; Score 65; DB 2; Length 459;
Best Local Similarity 75.0%; Pred. No. 0.0038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSLTHHNNIKPSSTR 16
DB 417 GPLTHHTNNKPSFTR 432

RESULT 8
T11372
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - donkey mitochondrion (fragment)
C/Species: Mitochondrion Equus asinus (donkey)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11372
R/Xu, X.; Gullberg, A.; Arnason, U.
J. Mol. Evol. 43, 438-463, 1996
A/Title: The complete mitochondrial (mtDNA) of the donkey and mtDNA comparisons among fo
A/Reference number: Z17265; MUID:97032591; PMID:8875857
A/Accession: T11372
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-459 <XUX>
A/Cross-references: UNIPROT:P2484; EMBL:X97337; NID:gl805746; PIDN:CAA66023.1; PID:gl80
A/Experimental source: kidney
C/Genetics:
A/Genome: mitochondrion
C/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 76.5%; Score 65; DB 2; Length 459;
Best Local Similarity 75.0%; Pred. No. 0.0038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSLTHHNNIKPSSTR 16
DB 417 GKYTHHNNIKPSFTR 432

RESULT 9
T11866
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - horse mitochondrion (fragment)
C/Species: Mitochondrion Equus caballus (domestic horse)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11866
```

```
R/Xu, X.; Arnason, U.
Gene 148, 357-362, 1994
A/Title: The complete mitochondrial DNA sequence of the horse, Equus caballus: Extensive
A/Reference number: Z17369; MUID:35047450; PMID:7958969
A/Accession: T11866
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-459 <XUX>
A/Cross-references: UNIPROT:P48655; EMBL:X79547; NID:G577571; PID:G577581; PIDN:CAA56081
C/Genetics:
A/Genome: mitochondrion
C/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 76.5%; Score 65; DB 2; Length 459;
Best Local Similarity 75.0%; Pred. No. 0.0038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSLTHHNNIKPSSTR 16
DB 417 GKYTHHNNIKPSFTR 432

RESULT 10
T11256
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - greater Indian rhinoceros mitoc
C/Species: Mitochondrion Rhinoceros unicornis (greater Indian rhinoceros)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11256
R/Xu, X.; Janke, A.; Arnason, U.
Mol. Biol. Evol. 13, 1167-1173, 1996
A/Title: The complete mitochondrial DNA sequence of the greater Indian rhinoceros, Rhin
A/Reference number: Z17256; MUID:97051708; PMID:8896369
A/Accession: T11256
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-459 <XUX>
A/Cross-references: UNIPROT:Q96068; EMBL:X97336; NID:gl666193; PIDN:CAA66010.1; PID:gl6
A/Experimental source: kidney
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 75.3%; Score 64; DB 2; Length 459;
Best Local Similarity 75.0%; Pred. No. 0.0057;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GSLTHHNNIKPSSTR 16
DB 417 GKYTHHNNIKPSFTR 432

RESULT 11
BS8851
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - finback whale mitochondrion
C/Species: Mitochondrion Balaenoptera physalus (finback whale, common rorqual)
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C/Accession: BS8851
R/Arnason, U.; Gullberg, A.; Widegren, B.
J. Mol. Evol. 33, 556-568, 1991
A/Title: The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Ba
A/Reference number: A58850; MUID:92139449; PMID:1779436
A/Accession: BS8851
A/Molecule type: DNA
A/Residues: 1-459 <ARN>
A/Cross-references: UNIPROT:P24975; GB:X61145; NID:gl2772; PIDN:CAA43448.1; PID:gl2780
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC1
C/Suprafamily: NADH dehydrogenase (ubiquinone) chain 4
```

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 72.9%; Score 62; DB 2; Length 459;

Best Local Similarity 75.0%; Pred. No. 0.012;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GSLTHHNNIKPSSTR 16

| | | | | | | | | | | | | |

Db 417 GKTHHNNITPSFTR 432

RESULT 12

Tl0981

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - pig mitochondrion

C;Species: mitochondrion Sus scrofa domestica (domestic pig)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: Tl0981; Tl1879

R;Lin, C.S.; Liu, C.X.; Sun, Y.L.; Chang, L.C.; Cheng, I.C.; Yang, P.C.; Mao, S.J.T.; Hu

submitted to the EMBL Data Library, November 1997

A;Description: Complete nucleotide sequence of the porcine mitochondrial genome.

A;Reference number: Z17237

A;Accession: Tl0981

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-459 <LIN>

A;Cross-references: UNIPROT:O79881; EMBL:AF034253; NID:g4958951; PIDN:AAD3

R;Ursing, B.M.

submitted to the EMBL Data Library, February 1999

A;Description: The complete mitochondrial DNA sequence of the pig (Sus scrofa).

A;Reference number: Z17370

A;Accession: Tl1879

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-360, 'V', 362-427, 'P', 429-459, 'C' <URS>

A;Cross-references: EMBL:AJ002189; PIDN:CAA05237.1

C;Genetics:

A;Gene: NADH4

A;Genome: mitochondrion

A;Genetic code: SGC1

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match

Best Local Similarity 72.9%; Score 62; DB 2; Length 459;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GSLTHHNNIKPSSTR 16

| | | | | | | | | | | | | |

Db 417 GKTHHNNIKASFTR 432

RESULT 13

Tl1059

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - sheep mitochondrion

C;Species: mitochondrion Ovis orientalis aries (domestic sheep)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: Tl1059

R;Hendler, S.; Lewalski, H.; Wassmuth, R.; Janke, A.

J. Mol. Evol. 47, 441-448, 1998

A;Title: The complete mitochondrial DNA sequence of the domestic sheep (Ovis aries) and

A;Reference number: Z17245; MUID:98440761; PMID:9767689

A;Accession: Tl1059

A;Status: preliminary;

A;Molecule type: DNA

A;Residues: 1-459 <HIE>

A;Cross-references: UNIPROT:O78755; EMBL:AF010406; NID:g3445513; PIDN:AAD1

A;Experimental source: Strain Merinolandschaf; liver

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

A;Note: ND4

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 71.8%; Score 61; DB 2; Length 459;

Best Local Similarity 75.0%; Pred. No. 0.018;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GSLTHHNNIKPSSTR 16

| | | | | | | | | | | | | |

Db 417 GKTHHNNILPSFTR 432

RESULT 14

Tl1502

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - dog mitochondrion (fragment)

C;Species: mitochondrion Canis lupus familiaris (dog)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: Tl1502

R;Kim, K.S.; Lee, S.E.; Jeong, H.W.; Ha, J.H.

Mol. Phylogenet. Evol. 10, 210-220, 1998

A;Title: The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochr

A;Reference number: Z17276; MUID:99097286; PMID:9878232

A;Accession: Tl1502

A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME

A;Molecule type: DNA

A;Residues: 1-459 <KIM>

A;Cross-references: UNIPROT:Q9Z258; EMBL:U96639; NID:g4154170; PID:g4154180; PIDN:RAD047

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

A;Note: ND4

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match

Best Local Similarity 70.6%; Score 60; DB 2; Length 459;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GSLTHHNNIKPSSTR 16

| | | | | | | | | | | | | |

Db 417 GYSHHNNIKPSFTR 432

Fri Nov 12 14:56:00 2004

Search completed: November 10, 2004, 15:55:06
Job time : 5.46619 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 26.8185 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-70

Perfect score: 85

Sequence: 1 GSUTHHNNKPSSTR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1925181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1925181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	89.4	152	2	AAF66138
2	76	89.4	459	1	NU4M_HUMAN
3	76	89.4	459	2	Q8RL39
4	76	89.4	459	2	Q8RQ99
5	76	89.4	459	2	Q8VHV8
6	76	89.4	459	2	Q8VI80
7	76	89.4	459	2	Q8VL53
8	76	89.4	459	2	Q8VLJ5
9	76	89.4	459	2	Q8WQ46
10	76	89.4	459	2	Q7Y780
11	76	89.4	459	2	Q7Y780
12	76	89.4	459	2	Q7Y891
13	76	89.4	459	2	Q7YCD8
14	76	89.4	459	2	Q7YCF7
15	76	89.4	459	2	Q7VEE1
16	76	89.4	459	2	Q85KV2
17	76	89.4	459	2	Q85KX4
18	76	89.4	459	2	Q85KY1
19	76	89.4	459	2	Q85KY4
20	76	89.4	459	2	Q85L09
21	76	89.4	459	2	Q8HB66
22	76	89.4	459	2	Q8HC47
23	76	89.4	459	2	Q8HG25
24	76	89.4	459	2	Q8HNR0
25	76	89.4	459	2	Q9T9W6
26	76	89.4	459	2	Q9T9Y1
27	76	89.4	459	2	Q9B078
28	76	89.4	459	2	Q9B188
29	76	89.4	459	2	Q9B1S7
30	76	89.4	459	2	Q9B2W2
31	76	89.4	459	2	Q9B2W5

```

32 76 89.4 459 2 Q9B2X6 Q9b2x6 homo sapien
33 76 89.4 459 2 Q9B2X8 Q9b2x8 homo sapien
34 76 89.4 459 2 Q9B2Y5 Q9b2y5 homo sapien
35 76 89.4 459 2 Q9B2Y7 Q9b2y7 homo sapien
36 76 89.4 459 2 Q9B300 Q9b300 homo sapien
37 76 89.4 459 2 Q8W8T1 Q8w8t1 homo sapien
38 76 89.4 459 2 Q8WCY0 Q8wcyc0 homo sapien
39 76 89.4 459 2 AAL54419 Aal54419 homo sapi
40 76 89.4 459 2 AAL54432 Aal54432 homo sapi
41 76 89.4 459 2 AAL54445 Aal54445 homo sapi
42 76 89.4 459 2 AAL54458 Aal54458 homo sapi
43 76 89.4 459 2 AAL54471 Aal54471 homo sapi
44 76 89.4 459 2 AAL54484 Aal54484 homo sapi
45 76 89.4 459 2 AAL54523 Aal54523 homo sapi

```

ALIGNMENTS

```

RESULT 1
AAF66138
ID AAF66138 PRELIMINARY; PRT; 152 AA.
AC AAF66138;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE URF4 GENE, PARTIAL CDS; TRNA-HIS, TRNA-SER, AND
DE TRNA-LEU GENES, COMPLETE SEQUENCE; AND URF5 GENE, PARTIAL CDS;
DE DE MITOCHONDRIAL GENES FOR MITOCHONDRIAL PRODUCTS (Fragment).
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82242101; PubMed=6284948;
RA Brown W.M., Prager E.M., Wang A., Wilson A.C.;
RT "Mitochondrial DNA sequences of primates: Tempo and mode of
RT evolution.";
RL J. Mol. Evol. 18:225-239(1982).
DR EMBL; L00016; AAF66138.1; -.
FT NON_TER 1
FT NON_TER 152
SQ SEQUENCE 152 AA; 16905 MW; E1A7D25096413C89 CRC64;

Query Match 89.4%; Score 76; DB 2; Length 152;
Best Local Similarity 87.5%; Pred.No. 0.00011;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSUTHHNNKPSSTR 16
Db 110 GSUTHHNNKPSSTR 125

RESULT 2
NU4M_HUMAN STANDARD; PRT; 459 AA.
ID NU4M_HUMAN
AC P03905; Q8HNR8;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN Name=MTND4; Synonyms=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81173052; PubMed=7219534;
RA Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,

```

RA Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
RA Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
RA "Sequence and organization of the human mitochondrial genome.";
RA Nature 290:457-465(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315217; PubMed=1377597;
RA Lu X., Walker T., Macmanus J.P., Seligy V.L.;
RA "Differentiation of HT-29 human colonic adenocarcinoma cells
RT correlates with increased expression of mitochondrial RNA: effects of
RT trehalose on cell growth and maturation.";
RA Cancer Res. 52:3718-3725(1992).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT PRO-79.
RX MEDLINE=22062553; PubMed=12022039;
RA Silva W.A. Jr., Bonatto S.L., Holanda A.J., Ribeiro-Dos-Santos A.K.,
RA Paixao B.M., Goldman G.H., Abe-Saunders K., Rodriguez-Delfin L.,
RA Barbosa M., Páco-Larson M.L., Petzl-Brler M.L., Valente V.,
RA Santos S.E., Zago M.A.;
RA "Mitochondrial genome diversity of native Americans supports a single
RT early entry of founder populations into America.";
RA Am. J. Hum. Genet. 71:187-192(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=12949126; DOI=10.1093/molbev/msg230;
RA Moilanen J.S., Finnila S., Majamaa K.;
RA "Lineage-specific selection in human mtDNA: lack of polymorphisms in a
RT segment of MTND5 gene in haplogroup J.";
RA Mol. Biol. Evol. 20:2132-2142(2003).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=21012010; PubMed=11130070; DOI=10.1038/35047064;
RA Ingman M., Kaessmann H., Paabo S., Gyllenstein U.;
RA "Mitochondrial genome variation and the origin of modern humans.";
RA Nature 408:708-713(2000).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=22723755; PubMed=12840039; DOI=10.1101/gr.686603;
RA Ingman M., Gyllenstein U.;
RA "Mitochondrial genome variation and evolutionary history of Australian
RT and New Guinean aborigines.";
RA Genome Res. 13:1600-1606(2003).
RN [7]
RP SEQUENCE FROM N.A.
RX PubMed=14760490; DOI=10.1007/s00414-004-0427-6;
RA Coble M.D., Just R.S., O'Callaghan J.E., Letmanyi I.H., Peterson C.T.,
RA Irwin J.A., Parsons T.J.;
RA "Single nucleotide polymorphisms over the entire mtDNA genome that
RT increase the power of forensic testing in Caucasians.";
RA Int. J. Legal Med. 118:137-146(2004).
RN [8]
RP SEQUENCE OF 308-459 FROM N.A.
RX MEDLINE=82242101; PubMed=6284948;
RA Brown W.M., Prager E.M., Wang A., Wilson A.C.;
RA "Mitochondrial DNA sequences of primates: tempo and mode of
RT evolution.";
RA J. Mol. Evol. 18:225-239(1982).
RN [9]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=85188293; PubMed=321850;
RA Chomyn A., Mariotti P., Clester M.W.J., Ragan C.I., Matsuno-Yagi A.,
RA Hatfi Y., Doolittle R.P., Atkadi G.;
RA "Six unidentified reading frames of human mitochondrial DNA encode
RT components of the respiratory-chain NADH dehydrogenase.";
RA Nature 314:592-597(1985).
RN [10]
RP VARIANT LHON HIS-340.
RX MEDLINE=89072713; PubMed=3201231;
RA Wallace D.C., Singh G., Lott M.T., Hodge J.A., Schurr T.G.,
RA Lezza A.M., Elsas L.J. II, Nikoskelainen E.K.;
RA "Mitochondrial DNA mutation associated with Leber's hereditary optic
RT neuropathy.";
RA Science 242:1427-1430(1988).
RN [11]
RP CHARACTERIZATION OF VARIANT LHON HIS-340.
RX MEDLINE=92070510; PubMed=1959619;
RA Majander A., Huoponen K., Savontaus M.-L., Nikoskelainen E.,
RA Wikstrom M.;
RA "Electron transfer properties of NADH:ubiquinone reductase in the
RT ND1/3460 and the ND4/11778 mutations of the Leber hereditary optic
RT neuroretinopathy (LHON).";
RA FEBS Lett. 292:289-292(1991).
RN [12]
RP VARIANT LHON HIS-340.
RX MEDLINE=92070953; PubMed=1959931;
RA Korman B.A., Schuster H., Berninger T.A., Leo-Kottler B.;
RA "Detection of the G to A mitochondrial DNA mutation at position 11778
RT in German families with Leber's hereditary optic neuropathy.";
RA Hum. Genet. 88:98-100(1991).
RN [13]
RP VARIANTS PRO-79; PRO-109; THR-132 AND THR-294.
RX MEDLINE=92098084; PubMed=1757091;
RA Marzuki S., Noer A.S., Lertit P., Thyagarajan D., Kapsa R.,
RA Uthmanaphol P., Byrne E.;
RA "Normal variants of human mitochondrial DNA and translation products:
RT the building of a reference data base.";
RA Hum. Genet. 88:139-145(1991).
RN [14]
RP VARIANT MELAS ALA-109.
RX MEDLINE=92359093; PubMed=1323207;
RA Lertit P., Noer A.S., Jean-Francois M.J.B., Kapsa R., Dennett X.,
RA Thyagarajan D., Lethlean K., Byrne E., Marzuki S.;
RA "A new disease-related mutation for mitochondrial encephalopathy
RT lactic acidosis and stroke-like episodes (MELAS) syndrome affects the
RT ND4 subunit of the respiratory complex I.";
RA Am. J. Hum. Genet. 51:457-468(1992).
RN [15]
RP VARIANT LDYT ILE-313.
RX MEDLINE=96220221; PubMed=8644732;
RA de Vries D.D., Went L.N., Bruyn G.W., Scholte H.R., Hofstra R.M.W.,
RA Bolhuis P.A., van Cost B.A.;
RA "Genetic and biochemical impairment of mitochondrial complex I
RT activity in a family with Leber hereditary optic neuropathy and
RT hereditary spastic dystonia.";
RA Am. J. Hum. Genet. 58:703-711(1996).
RN [16]
RP VARIANT LHON HIS-340.
RX MEDLINE=98112486; PubMed=9452107;
RA Sudoyo H., Sitepu M., Malik S., Poesponegoro H.D., Marzuki S.;
RA "Leber's hereditary optic neuropathy in Indonesia: two families with
RT the mtDNA 11778G>A and 14484T>C mutations.";
RA Hum. Mutat. Suppl. 1:S271-S274(1998).
RN [17]
RP -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- DISEASE: Defects in MTND4 are a cause of Leber's hereditary optic
CC neuropathy (LHON) [MIM:535000]; also known as Leber's optic
CC atrophy. LHON is a maternally inherited disease resulting in acute
CC bilateral blindness due to retinal degeneration predominantly in
CC young men. Cardiac conduction defects and neurological defects
CC have also been described, resulting in optic nerve degeneration
CC and cardiac dysrhythmia.
CC -1- DISEASE: Defects in MTND4 are a cause of Leber optic atrophy and
CC dystonia (LDYT) [MIM:500001]; also called familial dystonia with
CC visual failure and striatal lucencies.
CC -1- DISEASE: Defects in MTND4 are a cause of mitochondrial
CC encephalomyopathy with lactic acidosis and stroke-like episodes
CC syndrome (MELAS) [MIM:540000]. MELAS is a genetically
CC heterogeneous disorder, characterized by episodic vomiting,
CC seizures, and recurrent cerebral insults resembling strokes and
CC causing hemiparesis, hemianopsia, or cortical blindness.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@lsb-sib.ch).

CC EMBL; J01415; AAB58952.1; -

DR EMBL; V00662; CAA24035.1; -

DR EMBL; AY495090; AAR92503.1; -

DR EMBL; AY495091; AAR92516.1; -

DR EMBL; AY495092; AAR92529.1; -

DR EMBL; AY495093; AAR92542.1; -

DR EMBL; AY495094; AAR92555.1; -

DR EMBL; AY495095; AAR92568.1; -

DR EMBL; AY495096; AAR92581.1; -

DR EMBL; AY495097; AAR92594.1; -

DR EMBL; AY495098; AAR92607.1; -

DR EMBL; AY495099; AAR92620.1; -

DR EMBL; AY495100; AAR92633.1; -

DR EMBL; AY495101; AAR92646.1; -

DR EMBL; AY495102; AAR92659.1; -

DR EMBL; AY495103; AAR92672.1; -

DR EMBL; AY495104; AAR92685.1; -

DR EMBL; AY495105; AAR92698.1; -

DR EMBL; AY495106; AAR92711.1; -

DR EMBL; AY495107; AAR92724.1; -

DR EMBL; AY495108; AAR92737.1; -

DR EMBL; AY495109; AAR92750.1; -

DR EMBL; AY495110; AAR92763.1; -

DR EMBL; AY495111; AAR92776.1; -

DR EMBL; AY495112; AAR92789.1; -

DR EMBL; AY495113; AAR92802.1; -

DR EMBL; AY495114; AAR92815.1; -

DR EMBL; AY495115; AAR92828.1; -

DR EMBL; AY495116; AAR92841.1; -

DR EMBL; AY495117; AAR92854.1; -

DR EMBL; AY495118; AAR92867.1; -

DR EMBL; AY495119; AAR92880.1; -

DR EMBL; AY495120; AAR92893.1; -

DR EMBL; AY495121; AAR92906.1; -

DR EMBL; AY495122; AAR92919.1; -

DR EMBL; AY495123; AAR92932.1; -

Query Match 89.4%; Score 76; DB 1; Length 459;

Best Local Similarity 87.5%; Pred. No. 0.00036;

Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16

Db 417 GSLTHHNNKPSFTR 432

RESULT 3

Q6RL39 PRELIMINARY; PRT; 459 AA.

AC Q6RL39;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE NADH dehydrogenase subunit 4.

GN Name=ND4;

OS Homo sapiens (Human).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=14760490;

RA Coble M.D., Just R.S., O'Callaghan J.E., Letmanyi I.H., Peterson C.T.,

RA Irwin J.A., Parsons T.J.;

RT "Single nucleotide polymorphisms over the entire mtDNA genome that

increase the power of forensic testing in Caucasians.";

RL Int. J. Legal Med. 0:0-0(2004).

CC -/- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC EMBL; AY495316; AAR95443.1; -

DR EMBL; AY495316; AAR95443.1; -

DR GO; GO:0005739; C:mitochondrion; IEA.

DR InterPro; IPR003918; NADHub oxred4.

DR InterPro; IPR010227; NDH I_M.

DR InterPro; IPR001750; Oxidored_q1.

DR InterPro; IPR000360; Oxidored_q5_N.

DR Pfam; PF00361; Oxidored_q1; 1.

DR Pfam; PF01059; Oxidored_q5_N; 1.

DR PRINTS; PR01437; NUOXDRDTASE4.

DR TIGRFAMS; TIGR01972; NDH I_M; 1.

KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.

SQ SEQUENCE 459 AA; 51568 MW; 864A4A13C61244C7 CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;

Best Local Similarity 87.5%; Pred. No. 0.00036;

Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16

Db 417 GSLTHHNNKPSFTR 432

RESULT 4

Q6RQN9 PRELIMINARY; PRT; 459 AA.

AC Q6RQN9;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE NADH dehydrogenase subunit 4.

GN Name=ND4;

OS Homo sapiens (Human).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=14760490;

RA Coble M.D., Just R.S., O'Callaghan J.E., Letmanyi I.H., Peterson C.T.,

RA Irwin J.A., Parsons T.J.;

RT "Single nucleotide polymorphisms over the entire mtDNA genome that

increase the power of forensic testing in Caucasians.";

RL Int. J. Legal Med. 0:0-0(2004).

CC -/- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC EMBL; AY495147; AAR93246.1; -

DR EMBL; AY495147; AAR93246.1; -

DR GO; GO:0005739; C:mitochondrion; IEA.

DR InterPro; IPR003918; NADHub oxred4.

DR InterPro; IPR010227; NDH I_M.

DR InterPro; IPR001750; Oxidored_q1.

DR InterPro; IPR000260; Oxidored_q5_N.

DR Pfam; PF00361; Oxidored_q1; 1.

DR Pfam; PF01059; Oxidored_q5_N; 1.

DR PRINTS; PR01437; NUOXDRDTASE4.

DR TIGRFAMS; TIGR01972; NDH I_M; 1.

KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.

SQ SEQUENCE 459 AA; 51546 MW; FCBCFC59BDBD6971A CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;

Best Local Similarity 87.5%; Pred. No. 0.00036;

Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16

Db 417 GSLTHHNNKPSFTR 432

RESULT 5

Q6VHV8 PRELIMINARY; PRT; 459 AA.

ID Q6VHV8;

AC Q6VHV8;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE NADH dehydrogenase subunit 4.

```

GN Name=ND4;
OS Homo sapiens (Human).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313109; PubMed=11349229;
RA Finnilla S., Lehtonen M.S., Majamaa K.;
RT "Phylogenetic network for European mtDNA.";
RL Mol. Biol. Evol. 20:1475-1484(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12949126;
RA Moilanen J.S., Finnilla S., Majamaa K.;
RT "Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J.";
RL Mol. Biol. Evol. 20:2132-2142(2003).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=14760490;
RA Coble M.D., Just R.S., O'Callaghan J.E., Letmanyi I.H., Peterson C.T.,
IR Irwin J.A., Parsons T.J.;
RT "Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians.";
RL Int. J. Legal Med. 0:0-0(2004).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY339564; AAP91151.1; -
DR EMBL; AY339565; AAP91164.1; -
DR EMBL; AY495225; AAR94637.1; -
DR EMBL; AY339563; AAP91138.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PRO1437; NUOXDRDTASE4.
DR TIGRFAMs; TIGR01972; NDH I M; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51564 MW; 55C25A667D275BA CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHINNKPSSR 16
Db |||||:|||||
417 GSLTHINNKPSSR 432

RESULT 6
QY 1 GSLTHINNKPSSR 16
Db |||||:|||||
417 GSLTHINNKPSSR 432

RESULT 6
QY 1 GSLTHINNKPSSR 16
Db |||||:|||||
417 GSLTHINNKPSSR 432

ID Q6V180 PRELIMINARY; PRT; 459 AA.
AC Q6V180;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313109; PubMed=11349229;
RA Finnilla S., Lehtonen M.S., Majamaa K.;
RT "Phylogenetic network for European mtDNA.";
RL Mol. Biol. Evol. 20:1475-1484(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12949126;
RA Moilanen J.S., Finnilla S., Majamaa K.;
RT "Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J.";
RL Mol. Biol. Evol. 20:2132-2142(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY339430; AAP99409.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PRO1437; NUOXDRDTASE4.
DR TIGRFAMs; TIGR01972; NDH I M; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51580 MW; DBEE4D10917361B5 CRC64;

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12949126;
RA Moilanen J.S., Finnilla S., Majamaa K.;
RT "Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J.";
RL Mol. Biol. Evol. 20:2132-2142(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY339552; AAP90995.1; -
DR EMBL; AY339553; AAP91008.1; -
DR EMBL; AY339554; AAP90982.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PRO1437; NUOXDRDTASE4.
DR TIGRFAMs; TIGR01972; NDH I M; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51656 MW; 200E581F3BD270AF CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHINNKPSSR 16
Db |||||:|||||
417 GSLTHINNKPSSR 432

RESULT 7
QY 1 GSLTHINNKPSSR 16
Db |||||:|||||
417 GSLTHINNKPSSR 432

ID Q6VL53 PRELIMINARY; PRT; 459 AA.
AC Q6VL53;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313109; PubMed=11349229;
RA Finnilla S., Lehtonen M.S., Majamaa K.;
RT "Phylogenetic network for European mtDNA.";
RL Mol. Biol. Evol. 20:1475-1484(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12949126;
RA Moilanen J.S., Finnilla S., Majamaa K.;
RT "Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J.";
RL Mol. Biol. Evol. 20:2132-2142(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY339430; AAP99409.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR PRINTS; PRO1437; NUOXDRDTASE4.
DR TIGRFAMs; TIGR01972; NDH I M; 1.
DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51580 MW; DBEE4D10917361B5 CRC64;

```

Query Match 89.4%; Score 76; DB 2; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00036;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
 |||||:|||||
 Db 417 GSLTHHNNKPSSTR 432

RESULT 8
 ID Q6VLJ5 PRELIMINARY; PRT; 459 AA.
 AC Q6VLJ5;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE NADH dehydrogenase subunit 4.
 GN Name=ND4;
 OS Homo sapiens (Human).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2131109; PubMed=11349229;
 RA Finnila S., Lehtonen M.S., Majamaa K.;
 RT "Phylogenetic network for European mtDNA.";
 RL Am. J. Hum. Genet. 68:1475-1484(2001).
 RN [2]

SEQUENCE FROM N.A.
 RP PubMed=12949126;
 RA Moilanen J.S., Finnila S., Majamaa K.;
 RT "Lineage-specific selection in human mtDNA: lack of polymorphisms in a
 segment of MTND5 gene in haplogroup J.";
 RL Mol. Biol. Evol. 20:2132-2142(2003).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 DR EMEL; AY339404; AAP89071.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR010227; NDH I_M.
 DR InterPro; IPR001750; Oxidored_g1.
 DR InterPro; IPR00260; Oxidored_g5_N.
 DR Pfam; PF00361; Oxidored_g1; 1.
 DR Pfam; PF01059; Oxidored_g5_N; 1.
 DR PRINTS; PR01437; NUOXDRDTASE4.
 DR TIGRFAMs; TIGR01972; NDH I_M; 1.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 459 AA; 51652 MW; 2EE581C5BD270AF CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00036;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
 |||||:|||||
 Db 417 GSLTHHNNKPSSTR 432

RESULT 9
 ID Q6WQ46 PRELIMINARY; PRT; 459 AA.
 AC Q6WQ46;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE NADH dehydrogenase subunit 4.
 GN Name=ND4;
 OS Homo sapiens (Human).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

[1]
 RN SEQUENCE FROM N.A.
 RP PubMed=14563219;
 RA Maca-Meyer N., Gonzalez A.M., Pestano J., Flores C., Larruga J.M.,
 RA Cabrera V.M.;
 RT "Mitochondrial DNA transit between West Asia and North Africa inferred
 from U6 phylogeography.";
 RL BMC Genet. 4:15-15(2003).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 DR EMEL; AY275537; AAQ19469.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR010227; NDH I_M.
 DR InterPro; IPR001750; Oxidored_g1.
 DR InterPro; IPR00260; Oxidored_g5_N.
 DR Pfam; PF00361; Oxidored_g1; 1-
 DR Pfam; PF01059; Oxidored_g5_N; 1.
 DR PRINTS; PR01437; NUOXDRDTASE4.
 DR TIGRFAMs; TIGR01972; NDH I_M; 1.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 459 AA; 51640 MW; C79ESD1E37A37374 CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00036;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
 |||||:|||||
 Db 417 GSLTHHNNKPSSTR 432

RESULT 10
 ID Q7Y7B0 PRELIMINARY; PRT; 459 AA.
 AC Q7Y7B0;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE NADH dehydrogenase subunit 4.
 GN Name=ND4;
 OS Homo sapiens (Human).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22723755; PubMed=12840039;
 RA Ingman M., Gyllenstein U.;
 RT "Mitochondrial genome variation and evolutionary history of Australian
 and New Guinean aborigines.";
 RL Genome Res. 13:1600-1606(2003).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 DR EMEL; AY289057; AAP47967.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR010227; NDH I_M.
 DR InterPro; IPR001750; Oxidored_g1.
 DR InterPro; IPR00260; Oxidored_g5_N.
 DR Pfam; PF00361; Oxidored_g1; 1-
 DR Pfam; PF01059; Oxidored_g5_N; 1.
 DR TIGRFAMs; TIGR01972; NDH I_M; 1.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 459 AA; 51667 MW; 645F2183F908A55 CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
 Best Local Similarity 87.5%; Pred. No. 0.00036;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GSLTHHNNKPSSTR 16
DB 417 GSLTHHNNKPSFTR 432

RESULT 11
QY7S0
ID QY7S0 PRELIMINARY; PRT; 459 AA.
AC QY7S0;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DE 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22723755; PubMed=12840039;
RA Ingman M., Gyllenstein U.;
RT "Mitochondrial genome variation and evolutionary history of Australian
and New Guinean aborigines.";
RL Genome Res. 13:1600-1606(2003).

RESULT 12
QY7S0
ID QY7S0 PRELIMINARY; PRT; 459 AA.
AC QY7S0;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DE 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22723755; PubMed=12840039;
RA Ingman M., Gyllenstein U.;
RT "Mitochondrial genome variation and evolutionary history of Australian
and New Guinean aborigines.";
RL Genome Res. 13:1600-1606(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY289053; AAP47915.1; -.
DR EMBL; AY289072; AAP48162.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0042773; P:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR000260; Oxidored_g5_N.
DR Pfam; PF00361; Oxidored_g1; 1.
DR Pfam; PF01059; Oxidored_g5_N; 1.
DR TIGRFAWS; TIGR01972; NDH I M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51553 MW; 34DD711F3BC3915F CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
DB 417 GSLTHHNNKPSFTR 432

RESULT 12
QY7S0
ID QY7S0 PRELIMINARY; PRT; 459 AA.
AC QY7S0;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DE 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22723755; PubMed=12840039;
RA Ingman M., Gyllenstein U.;
RT "Mitochondrial genome variation and evolutionary history of Australian
and New Guinean aborigines.";
RL Genome Res. 13:1600-1606(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY289053; AAP47915.1; -.
DR EMBL; AY289072; AAP48162.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0042773; P:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR000260; Oxidored_g5_N.
DR Pfam; PF00361; Oxidored_g1; 1.
DR Pfam; PF01059; Oxidored_g5_N; 1.
DR TIGRFAWS; TIGR01972; NDH I M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51553 MW; 34DD711F3BC3915F CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
DB 417 GSLTHHNNKPSFTR 432

RESULT 13
QY7S0
ID QY7S0 PRELIMINARY; PRT; 459 AA.
AC QY7S0;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Kong Q.-P., Yao Y.-G., Sun C., Bandelt H.-J., Zhu C.-L., Zhang Y.-P.;
RT "Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from
Complete Sequences.";
RL Am. J. Hum. Genet. 0:0-0(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY255163; AAO67017.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0042773; P:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_g1.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22723755; PubMed=12840039;
RA Ingman M., Gyllenstein U.;
RT "Mitochondrial genome variation and evolutionary history of Australian
and New Guinean aborigines.";
RL Genome Res. 13:1600-1606(2003).

RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=14760490;
RA Coble M.D., Just R.S., O'Callaghan J.E., Letmanyi I.H., Peterson C.T.,
Irwin J.A., Parsons T.J.;
RT "Single nucleotide polymorphisms over the entire mtDNA genome that
increase the power of forensic testing in Caucasians.";
RL Int. J. Legal Med. 0:0-0(2004).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY255139; AAO66705.1; -.
DR EMBL; AY289066; AAP48084.1; -.
DR EMBL; AY289067; AAP48097.1; -.
DR EMBL; AY495166; AAR93493.1; -.
DR EMBL; AY495258; AAR94689.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0042773; P:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR000260; Oxidored_g5_N.
DR Pfam; PF00361; Oxidored_g1; 1.
DR Pfam; PF01059; Oxidored_g5_N; 1.
DR TIGRFAWS; TIGR01972; NDH I M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51608 MW; 02E1F27CCF797BF5 CRC64;

Query Match 89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHNNKPSSTR 16
DB 417 GSLTHHNNKPSFTR 432

RESULT 13
QY7S0
ID QY7S0 PRELIMINARY; PRT; 459 AA.
AC QY7S0;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Kong Q.-P., Yao Y.-G., Sun C., Bandelt H.-J., Zhu C.-L., Zhang Y.-P.;
RT "Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from
Complete Sequences.";
RL Am. J. Hum. Genet. 0:0-0(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY255163; AAO67017.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0042773; P:ATP synthetase coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH I M.
DR InterPro; IPR001750; Oxidored_g1.

```

```

DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR TIGRFAMs; TIGR01972; NDH_I_M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51606 MW; 564652E36FC7851B CRC64;

Query Match      89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHINNKPSTR 16
   |||||:||||:||||
Db 417 GSLTHHINNKPSTR 432

RESULT 14
Q7YCF7 PRELIMINARY; PRT; 459 AA.
AC Q7YCF7;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kong O.-P., Yao Y.-G., Sun C., Bandelt H.-J., Zhu C.-L., Zhang Y.-P.;
RT "Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from
RT Complete Sequences.";
RL Am. J. Hum. Genet. 0:0-0(2003).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY255147; AAO66809.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH_I_M.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR TIGRFAMs; TIGR01972; NDH_I_M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51640 MW; 8444F6A48FC362AE CRC64;

Query Match      89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHINNKPSTR 16
   |||||:||||:||||
Db 417 GSLTHHINNKPSTR 432

RESULT 15
Q7YEE1 PRELIMINARY; PRT; 459 AA.
AC Q7YEE1;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ND4;
OS Homo sapiens (Human).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ingman M., Gyllenstein U.;
RT "Mitochondrial genome variation and evolutionary history of Australian
RT and New Guinean aborigines.";
RL Genome Res. 13:1600-1606(2003).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; AY289086; AAP48344.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010227; NDH_I_M.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR00260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
DR TIGRFAMs; TIGR01972; NDH_I_M; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 459 AA; 51550 MW; 2EE5805917361AF CRC64;

Query Match      89.4%; Score 76; DB 2; Length 459;
Best Local Similarity 87.5%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSLTHHINNKPSTR 16
   |||||:||||:||||
Db 417 GSLTHHINNKPSTR 432

Search completed: November 10, 2004, 15:53:33
JOB time : 27.8954 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 12.9537 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-71

Perfect score: 141

Sequence: 1 VSCWPSYLKYLSTASASLLATQLKS 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	34.4	372	4	US-09-252-991A-32717
2	48	34.0	50	4	US-08-469-260A-214
3	48	34.0	50	4	US-08-488-446-214
4	48	34.0	50	4	US-08-467-344A-214
5	48	34.0	50	4	US-08-424-550B-214
6	48	34.0	1244	3	US-08-938-291A-5
7	48	34.0	1244	4	US-09-589-619-5
8	47	33.3	368	4	US-09-252-991A-21362
9	46	32.6	235	4	US-09-252-991A-18300
10	46	32.6	337	3	US-09-032-372-2
11	46	32.6	409	4	US-08-248-796A-21624
12	46	32.6	438	4	US-09-252-991A-23034
13	46	32.6	594	4	US-08-938-291A-6
14	46	32.6	1349	3	US-08-589-619-6
15	46	32.6	1349	4	US-09-589-619-6
16	45.5	32.3	581	3	US-08-619-812-6
17	45	31.9	99	4	US-09-270-767-31886
18	45	31.9	99	4	US-09-270-767-47103
19	45	31.9	193	4	US-09-270-767-32917
20	45	31.9	193	4	US-08-270-767-48134
21	45	31.9	197	4	US-08-248-796A-15582
22	45	31.9	235	4	US-09-248-796A-14353
23	45	31.9	245	4	US-09-270-767-35096
24	45	31.9	245	4	US-09-270-767-50313
25	45	31.9	489	5	PCT-US95-11405-35
26	44	31.2	100	4	US-08-270-767-34940
27	44	31.2	100	4	US-09-270-767-50157

28 44 31.2 118 4 US-09-248-796A-26791 Sequence 26791, A
29 44 31.2 126 4 US-09-270-767-38734 Sequence 38734, A
30 44 31.2 126 4 US-09-270-767-53951 Sequence 53951, A
31 44 31.2 390 4 US-09-107-532A-5233 Sequence 5233, Ap
32 44 31.2 512 4 US-09-543-681A-4510 Sequence 4510, Ap
33 44 31.2 570 3 US-08-832-078-5 Sequence 5, Appli
34 44 31.2 787 4 US-09-698-789B-5 Sequence 31, Appl
35 44 31.2 806 4 US-09-187-330-31 Sequence 3, Appli
36 44 31.2 828 4 US-09-187-330-55 Sequence 55, Appl
37 44 31.2 289 4 US-09-252-991A-27509 Sequence 27509, A
38 43.5 30.9 499 4 US-09-252-991A-17198 Sequence 17198, A
39 43.5 30.9 1614 3 US-09-052-469-2 Sequence 2, Appli
40 43.5 30.9 1614 4 US-08-422-582-2 Sequence 2, Appli
41 43.5 30.9 1614 4 US-09-052-262-2 Sequence 2, Appli
42 43.5 30.9 4302 3 US-08-658-136-5 Sequence 5, Appli
43 43.5 30.9 4302 3 US-09-052-489-8 Sequence 8, Appli
44 43.5 30.9 4302 4 US-08-422-582-8 Sequence 8, Appli
45 43.5 30.9 4302 4 US-08-422-582-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-252-991A-32717

; Sequence 32717, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 32717

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32717

Query Match 34.4%; Score 48.5; DB 4; Length 372;

Best Local Similarity 34.3%; Pred. No. 17;

Matches 12; Conservative 5; Mismatches 9; Indels 9; Gaps 1;

Qy 1 VSCWPS-----YLYKYLSTASASLLATQLKS 26

Db 336 VSCWFPADTHPRCSRRYGSSTPSATTAIRATAKA 370

RESULT 2

US-08-469-260A-214

; Sequence 214, Application US/08469260A

; Patent No. 6451578

; GENERAL INFORMATION:

; APPLICANT: JOHN N. SIMONS

; APPLICANT: TAMI J. PILOT-MATTIAS

; APPLICANT: GEORGE J. DAWSON

; APPLICANT: GEORGE G. SCHLAUDER

; APPLICANT: SURESH M. DESAI

; APPLICANT: THOMAS P. LEARY

; APPLICANT: ANTHONY SCOTT MUEHROFF

; APPLICANT: JAMES C. ERKER

; APPLICANT: SHERI L. BUIJK

; APPLICANT: ISA K. MUSHAWAR

; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS

; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE

; NUMBER OF SEQUENCES: 716

; CORRESPONDENCE ADDRESS:


```

; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 214:
US-08-467-344A-214

Query Match 34.0%; Score 48; DB 4; Length 50;
Best Local Similarity 40.0%; Pred. No. 1.7;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 4 WPSYLYPLSTASASLLATOLKSA 28
Db 9 WPAVSTPLCTSEAKLTFQLVTC 33

RESULT 5
US-06-424-550B-214
; Sequence 214, Application US/08424550B
; Patent No. 6720166
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/APED
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-214

Query Match 34.0%; Score 48; DB 4; Length 50;
Best Local Similarity 40.0%; Pred. No. 1.7;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 4 WPSYLYPLSTASASLLATOLKSA 28

```

```

Db 9 WPAVSTPLCTSEAKLTFQLVTC 33

RESULT 6
US-08-938-291A-5
; Sequence 5, Application US/08938291A
; Patent No. 6117673
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: RDGB PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291A
; FILING DATE: September 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,337
; FILING DATE: October 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 228/172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-938-291A-5

Query Match 34.0%; Score 48; DB 3; Length 1244;
Best Local Similarity 45.5%; Pred. No. 90;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 5 PSYLYPLSTASASLLATOLKS 26
Db 757 PRYQFPLGDSGLLLADTLQT 778

RESULT 7
US-09-589-619-5
; Sequence 5, Application US/09589619
; Patent No. 6576442
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: RDGB PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/589,619
; FILING DATE: 07-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291
; FILING DATE: September 26, 1997
; APPLICATION NUMBER: 60/027,337
; FILING DATE: October 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 228/172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-589-619-5
Query Match 34.0%; Score 48; DB 4; Length 1244;
Best Local Similarity 45.5%; Pred. No. 90;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 5 PSYKYLPLSTASASLLATOLKS 26
DB 757 PRYQKFLDGGSSLLATLQT 778
RESULT 8
US-09-252-991A-21362
; Sequence 21362, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21362
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21362
Query Match 33.3%; Score 47; DB 4; Length 368;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 12; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 3 CWPSYLYKYLPLSTASASLLATOL 24
DB 292 CMEAVLKRFPVRRREASFFSLGGDSLLATRL 321
RESULT 9
US-09-252-991A-18300
; Sequence 18300, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18300
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18300
Query Match 32.6%; Score 46; DB 4; Length 235;
Best Local Similarity 47.1%; Pred. No. 24;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 2 SCWPSYLYKYLPLSTASAS 18
DB 28 AAWPSTWTPSPASTANAT 44
RESULT 10
US-09-032-372-2
; Sequence 2, Application US/09032372
; Patent No. 6008337
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,372
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0478 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRSN0202
; CLONE: 680517
US-09-032-372-2

Query Match 32.6%; Score 46; DB 3; Length 337;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 PSYLYPLSTASASL 19
DB 306 PDYLYSISTALCSL 320

RESULT 11

US-09-248-796A-21624
; Sequence 21624, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21624
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-21624

Query Match 32.6%; Score 46; DB 4; Length 409;
Best Local Similarity 35.0%; Pred. No. 48;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 5 PSYLYPLSTASASLTLQ 24
DB 268 PNYMPYLSSEKSTLTPSEV 287

RESULT 12

US-09-252-991A-23034
; Sequence 23034, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23034
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23034

Query Match 32.6%; Score 46; DB 4; Length 438;
Best Local Similarity 33.3%; Pred. No. 52;
Matches 7; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKPLSTASASLLA 21
DB 39 IARWPSRWRPISSSSPAAMA 59

RESULT 13

US-09-252-991A-23804
; Sequence 23804, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23804
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23804

Query Match 32.6%; Score 46; DB 4; Length 594;
Best Local Similarity 45.5%; Pred. No. 75;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKPLSTASASLLAT 22
DB 33 VGCWRNTASSPTPTASSAPAT 54

RESULT 14

US-08-938-291A-6
; Sequence 6, Application US/08938291A
; Patent No. 6117673
; GENERAL INFORMATION:

; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDBG PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291A

Search completed: November 10, 2004, 15:57:22
Job time : 12.9537 secs

TELEFAX: (213) 67-3510

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 67.6584 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-71

Perfect score: 141

Sequence: 1 VSCWPSYLYKPLSTASASLLATQLKSIA 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	141	100.0	28	US-10-092-750-71	Sequence 71, Appl
2	57	40.4	928	US-10-104-047-2435	Sequence 2435, Ap
3	55	39.0	349	US-10-424-599-163030	Sequence 163030,
4	53	37.6	173	US-10-424-599-167104	Sequence 167104,
5	53	37.6	1257	US-10-109-324-2	Sequence 2, Appl1
6	52	36.9	56	US-10-424-599-143889	Sequence 143889,
7	51	36.2	116	US-09-764-891-4490	Sequence 4490, Ap
8	51	36.2	697	US-10-425-115-332665	Sequence 332665,
9	51	36.2	715	US-10-425-114-65245	Sequence 65245, A
10	50	35.5	176	US-10-437-963-132933	Sequence 132933,
11	49.5	35.1	76	US-10-424-599-279483	Sequence 279483,
12	49	34.8	84	US-10-425-115-304497	Sequence 304497,
13	48.5	34.4	113	US-10-108-260A-3204	Sequence 3204, Ap

Sequence 214, App
Sequence 174891,
Sequence 145804,
Sequence 177818,
Sequence 266647,
Sequence 67673, A
Sequence 230760,
Sequence 201026,
Sequence 284922,
Sequence 28487,
Sequence 3936, Ap
Sequence 122676,
Sequence 105958,
Sequence 77616, A
Sequence 13391, A
Sequence 30, Appl
Sequence 83, Appl
Sequence 342899,
Sequence 838, App
Sequence 361069,
Sequence 170620,
Sequence 18726,
Sequence 145753,
Sequence 4118, Ap
Sequence 208, App
Sequence 22, Appl
Sequence 23, Appl
Sequence 119, App
Sequence 8789, Ap
Sequence 42, Appl
Sequence 2341, Ap

8 US-08-424-550B-214
15 US-10-424-599-174891
16 US-10-437-963-145804
17 US-10-437-963-177818
18 US-10-425-115-266647
19 US-10-282-122A-67673
20 US-10-425-115-230760
21 US-10-437-963-201026
22 US-10-425-115-284922
23 US-10-424-599-28487
24 US-10-108-260A-3936
25 US-10-437-963-105958
26 US-10-437-963-122676
27 US-10-282-122A-77616
28 US-10-389-493-13391
29 US-10-324-967-30
30 US-10-408-765A-83
31 US-10-425-115-342899
32 US-10-083-357-838
33 US-10-425-115-361069
34 US-10-424-599-170620
35 US-09-784-891-3638
36 US-10-437-963-118726
37 US-10-437-963-145753
38 US-09-738-626-4118
39 US-10-781-014-208
40 US-10-463-782A-22
41 US-10-463-782A-23
42 US-10-108-605-119
43 US-10-369-493-8789
44 US-09-847-102A-42
45 US-10-408-765A-2341

ALIGNMENTS

RESULT 1
US-10-092-750-71
; Sequence 71, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-71

Query Match 100.0%; Score 141; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSCWPSYLYKPLSTASASLLATQLKSIA 28
Db 1 VSCWPSYLYKPLSTASASLLATQLKSIA 28

RESULT 2
US-10-104-047-2435
; Sequence 2435, Application US/10104047
; Publication No. US2003023692A1
; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2435
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2435

Query Match 40.4%; Score 57; DB 14; Length 928;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKYPPLSTASASLL 20
: ||| ||| ||| :
Db 241 IECWPSNSKYPLVPVFTFVL 260

RESULT 3

US-10-424-599-163030
; Sequence 163030, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163030
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118233C.1.pep
US-10-424-599-163030

Query Match 39.0%; Score 55; DB 15; Length 349;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 5 PSYLYPLSTASASLLATQLKS 26
||| : ||| : ||| :
Db 214 PLYLRPVMTAEPSYLARKLKA 235

RESULT 4

US-10-424-599-167104
; Sequence 167104, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167104
; LENGTH: 173

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(173)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121909C.1.pep
US-10-424-599-167104

Query Match 37.6%; Score 53; DB 15; Length 173;
Best Local Similarity 57.1%; Pred. No. 10;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 SYLYKPLSTASASLLATQLKS 26
|:| ||| ||| ||| :
Db 87 SFUTRPLSTLSUSRLASQIES 107

RESULT 5

US-10-109-324-2
; Sequence 2, Application US/10109324
; Publication No. US20030082682A1
; GENERAL INFORMATION:
; APPLICANT: Hinayana L. Bawagan
; APPLICANT: Kathryn B. Freeman
; APPLICANT: Xiatong Li
; TITLE OF INVENTION: Polynucleotide and Polypeptide Sequence
; FILE REFERENCE: of RXR Interacting Protein, RB204
; CURRENT APPLICATION NUMBER: US/10/109,324
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,397
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1257
; TYPE: PRT
; ORGANISM: human
US-10-109-324-2

Query Match 37.6%; Score 53; DB 14; Length 1257;
Best Local Similarity 40.0%; Pred. No. 93;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKYPPLSTASASLL 20
: ||| : ||| : ||| :
Db 5 IECWPAQTKYPPVFTFVL 24

RESULT 6

US-10-424-599-143889
; Sequence 143889, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143889
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100945C.1.pep
US-10-424-599-143889


```
Query Match      36.9%; Score 52; DB 15; Length 56;
Best Local Similarity 46.2%; Pred. No. 4.3;
Matches 12; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 SCWPSYLYKYPSTASASLLATQKSI 27
DB 25 SSTPSFIKTPLASFALSITQOLVHI 50

RESULT 7
US-09-764-891-4490
; Sequence 4490, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4490
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4490

Query Match      36.2%; Score 51; DB 10; Length 116;
Best Local Similarity 34.8%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKYPSTASASLLATQ 23
DB 93 LECWPLMTSHPFPGPSLMTLAVE 115

RESULT 8
US-10-425-115-332665
; Sequence 332665, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 332665
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_66505C.1.pep
US-10-425-115-332665

Query Match      36.2%; Score 51; DB 17; Length 697;
Best Local Similarity 53.8%; Pred. No. 96;
Matches 14; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 5 PSYLYKYPSTASAS--LLATQKSLA 28
DB 18 PSYLYPLCTDARAPVLLNTTLKFA 43
```

```
RESULT 9
US-10-425-114-65245
; Sequence 65245, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65245
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4745-009-A8_FLI.pep
US-10-425-114-65245

Query Match      36.2%; Score 51; DB 15; Length 715;
Best Local Similarity 53.8%; Pred. No. 99;
Matches 14; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 5 PSYLYKYPSTASAS--LLATQKSLA 28
DB 36 PSYLYPLCTDARAPVLLNTTLKFA 61

RESULT 10
US-10-437-963-132933
; Sequence 132933, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132933
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_34854C.1.pep
US-10-437-963-132933

Query Match      35.5%; Score 50; DB 16; Length 176;
Best Local Similarity 47.6%; Pred. No. 30;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 SCWPSYLYKYPSTASASLLAT 22
DB 127 SCFPHSVKRSLSSSSSSQGT 147

RESULT 11
US-10-424-599-279483
```

```

; Sequence 279483, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 279483
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_94396C.1.pap
US-10-424-599-279483

Query Match      35.1%; Score 49.5; DB 15; Length 76;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY      4 WPSYLYPLSTAS-ASLLATQK 25
      ||:|||||:||||:|
DB      47 WPSLYLSCVSVASILSPHIQ 69
      ||:|||||:||||:|

RESULT 12
US-10-425-115-304497
; Sequence 304497, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 304497
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_40771C.1.pap
US-10-425-115-304497

Query Match      34.8%; Score 49; DB 17; Length 84;
Best Local Similarity 40.9%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY      4 WPSYLYPLSTASASLLATQK 25
      |||:|:|:|:|:|:|
DB      61 WPSWTFPFSTVVERLYLELK 82
      |||:|:|:|:|:|:|

RESULT 13
US-10-108-260A-3204
; Sequence 3204, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3204
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3204

Query Match      34.4%; Score 48.5; DB 15; Length 113;
Best Local Similarity 61.1%; Pred. No. 31;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY      2 SCWPSYLYKYP-ISTASAS 18
      |||||:|:|:|:|
DB      75 SCWPSWSQTPDLSDPSAS 92
      |||||:|:|:|:|

RESULT 14
US-08-424-550B-214
; Sequence 214, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-938-3635
; TELEFAX: 708-938-3623
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-214

Query Match      34.0%; Score 48; DB 8; Length 50;
Best Local Similarity 40.0%; Pred. No. 15;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY      4 WPSYLYPLSTASASLLATQK 28
      ||:|:|:|:|:|:|
DB      9 WPAVSTPLCTSEAKKLTFLVATCA 33
      ||:|:|:|:|:|:|

```


This Page Blank (uspio)

[illegible]

RESULT 12

C95847
 hypothetical transmembrane protein SMB20040 [imported] - Sinorhizobium meliloti (strain C) Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C/Accession: C95847
 R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A/Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo A/Reference number: A95842; MUID:21396508; PMID:11481431
 A/Accession: C95847
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-286 <KUR>
 A/Cross-references: UNIPROT:Q92XB9; GB:AL591985; PIDN:CAC48443.1; PID:gl5139915; GSPDB:C A/Experimental source: strain 1021, megaplasmid pSymb
 R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, R.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A/Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation
 C/Genetics:
 A/Gene: SMB20040
 A/Genome: plasmid

Query Match 32.6%; Score 46; DB 2; Length 286;
 Best Local Similarity 42.9%; Pred. No. 37;
 Matches 12; Conservative 3; Mismatches 11; Indels 2; Gaps 1;

QY 1 VSCWPSYLKYLPLSTASLLATQLKSLA 28

DB 179 VICNALVLSLPVMT-ALLMLATMPSSLA 204

RESULT 13

C97013
 probable non-processive endoglucanase family 5, ortholog of mannase A, secreted, docked C) Species: Clostridium acetobutylicum
 C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C/Accession: C97013
 R/Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R. J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: C97013
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-425 <KUR>
 A/Cross-references: UNIPROT:Q97XK1; GB:AE001437; PIDN:AAK78894.1; PID:gl5023818; GSPDB:C A/Experimental source: Clostridium acetobutylicum AFCC924
 C/Genetics:
 A/Gene: CAC0918

Query Match 32.6%; Score 46; DB 2; Length 425;
 Best Local Similarity 52.6%; Pred. No. 57;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLKYLPLSTASLL 19

DB 229 VDCIHQYAKQFVSGSASL 247

RESULT 14

S52677
 probable membrane protein YDR111c - yeast (Saccharomyces cerevisiae)
 N/Alternate names: hypothetical protein YD9727.07c
 C) Species: Saccharomyces cerevisiae

C>Date: 19-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C/Accession: S52677
 R/Murphy, L.; Shore, L.; Harris, D. submitted to the EMBL Data Library, March 1995
 A/Reference number: S52671
 A/Accession: S52677
 A/Molecule type: DNA
 A/Residues: 1-507 <MUR>
 A/Cross-references: UNIPROT:P52892; EMBL:Z48758; NID:g747879; PIDN:CAA88665.1; PID:g7478

C/Genetics:
 A/Gene: MIPS.YDR111C
 A/Cross-references: SGD:S0002518
 A/Map position: 4R
 C/Superfamily: alanine transaminase
 C/Keywords: transmembrane protein
 F;265-281/Domain: transmembrane #status predicted <TM1>
 F;359-375/Domain: transmembrane #status predicted <TM2>

Query Match 32.6%; Score 46; DB 2; Length 507;
 Best Local Similarity 62.5%; Pred. No. 69;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 KYPLSTASASLLATQL 24

DB 198 QYPLTASASLNFNAQV 213

RESULT 15

S03540
 gene frizzled protein precursor - fruit fly (Drosophila melanogaster) C) Species: Drosophila melanogaster
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
 C/Accession: S03540; S15708; S15709
 R/Vinson, C.R.; Conover, S.; Adler, P.N. Nature 338, 263-264, 1989
 A/Title: A Drosophila tissue polarity locus encodes a protein containing seven potentia A/Reference number: S03540; MUID:89159415; PMID:2493583
 A/Accession: S03540
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-581 <VIN>

A/Cross-references: UNIPROT:P18337
 R/Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L. Genetics 126, 401-416, 1990
 A/Title: Molecular structure of frizzled, a Drosophila tissue polarity gene.
 A/Reference number: S15708; MUID:91060073; PMID:2174014
 A/Accession: S15708
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-581 <ADL>
 A/Cross-references: EMBL:X54648; NID:g7973; PIDN:CAA38460.1; PID:g804979
 A/Accession: S15709
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-405, 'MY', 408, 'WQFHTIN' <AD2>
 A/Cross-references: EMBL:X54648; NID:g7973; PIDN:CAA38461.1; PID:g804980
 C/Genetics:

A/Gene: FlyBase:fbz
 A/Cross-references: FlyBase:FBgn0001085
 A/Introns: 224/3; 264/3; 329/3; 405/1
 C/Superfamily: fruit fly frizzled protein
 C/Keywords: alternative splicing; transmembrane protein
 F;1-26/Domain: signal sequence #status predicted <SIG>
 F;27-581/Product: gene frizzled protein #status predicted <MAP>

Query Match 32.6%; Score 46; DB 2; Length 581;
 Best Local Similarity 33.3%; Pred. No. 79;
 Matches 13; Conservative 6; Mismatches 6; Indels 14; Gaps 2;

QY 4 WFSYL---KYPF-----STASASLLATQLKSLA 28

DB 145 WPNLECKFPVHGEGDLCVAENTTSSASTAATPTPSVA 183

Search completed: November 10, 2004, 15:55:07
Job time : 10.5658 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 46.9324 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-71
Perfect score: 141
Sequence: 1 VSCWPSYLKYPFLSTASILLATQLKSTIA 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	40.4	928	Q8NB93	Q8NB93 homo sapien
2	57	40.4	1831	Q86T77	Q86T77 homo sapien
3	57	40.4	1863	Q8IVX2	Q8IVX2 homo sapien
4	57	40.4	1865	Q7Z401	Q7Z401 homo sapien
5	53	37.6	1369	Q6NZJ7	Q6NZJ7 mus muscu
6	53	37.6	1369	AAH66102	AAH66102 mus muscu
7	51.5	36.5	118	Q6GZ70	Q6GZ70 homo sapien
8	51.5	36.5	147	Q6GZ72	Q6GZ72 homo sapien
9	51	36.2	306	Q8ZJ53	Q8ZJ53 yersinia pe
10	51	36.2	320	Q8D132	Q8D132 yersinia pe
11	51	36.2	320	AA560695	AA560695 yersinia
12	51	36.2	394	Q9JGK4	Q9JGK4 meleagrid h
13	50	35.5	565	P73143	P73143 synchocyst
14	50	35.5	608	Q7X329	Q7X329 uncultured
15	50	35.5	760	Q47500	Q47500 venturia in
16	49.5	35.1	916	Q7N3S1	Q7N3S1 photorhabdu
17	49	34.8	211	Q8ZVC8	Q8ZVC8 photobaculu
18	49	34.8	249	Q9ADM9	Q9ADM9 streptomyce
19	49	34.8	353	Q8XRK7	Q8XRK7 ralsstonia s
20	49	34.8	574	Q7SFI4	Q7SFI4 neosporea
21	49	34.8	850	Q6C8F7	Q6C8F7 yarrowia li
22	48.5	34.4	137	Q97X41	Q97X41 sulfolobus
23	48	34.0	120	Q9L224	Q9L224 streptomyce
24	48	34.0	240	Q9VKJ5	Q9VKJ5 drosophila
25	48	34.0	321	Q7NYI9	Q7NYI9 chromobacte
26	48	34.0	323	Q88CN9	Q88CN9 pseudomonas
27	48	34.0	378	Q8ITN3	Q8ITN3 ceratitidis c
28	48	34.0	630	1 S6A4	1 S6A4 CAVPO
29	48	34.0	686	2 Q8VYGO	2 Q8VYGO arabidopsis
30	48	34.0	978	2 Q8F235	2 Q8F235 arabidopsis
31	48	34.0	1008	2 Q6CX67	2 Q6CX67 kluyveromyc

32	48	34.0	1243	2	Q6T7X3	Q6T7X3 homo sapien
33	48	34.0	1243	2	Q8TBN3	Q8TBN3 homo sapien
34	48	34.0	1243	2	Q35954	Q35954 mus musculu
35	48	34.0	1243	2	AAR06909	AAR06909 homo sapi
36	48	34.0	1244	2	O00562	O00562 homo sapien
37	48	34.0	1244	2	Q8BZ73	Q8BZ73 homo sapien
38	48	34.0	1270	2	Q8VUL1	Q8VUL1 bacillus th
39	48	34.0	1299	2	Q847C8	Q847C8 nodularia s
40	47.5	33.7	180	2	Q6X100	Q6X100 drosophila
41	47.5	33.7	180	2	AAR10055	AAR10055 drosophila
42	47.5	33.7	190	2	Q8VL69	Q8VL69 drosophila
43	47.5	33.7	298	2	Q8RS30	Q8RS30 comamonas a
44	47.5	33.7	298	2	BAC1981	BAC1981 comamonas
45	47.5	33.7	345	2	Q8BLT8	Q8BLT8 mus musculu

ALIGNMENTS

RESULT 1
Q8NB93
ID Q8NB93 PRELIMINARY; PRT; 928 AA.
AC Q8NB93;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein FLJ34049.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039;
RA Wakanatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Shiratori A.,
Murakami K., Sasada T., Iwayanagi T., Wagatsuma M., Shiratori A.,
Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
Oono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Ohtsuka S.,
Musashino K., Yuuki H., Oshima A., Sasaki N., Arita M., Imose N.,
Yoshikawa S., Matsumawa H., Ichihara T., Shiohara N., Sano S.,
Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
Nakagawa S., Sanoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
Yamagaki H., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi R.,
Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs".
RL Nat. Genet. 36:40-45(2004).
RL ENBL; AK091368; BAC03648.1; -.
DR InterPro; IPR005112; dDenn.
DR InterPro; IPR001194; DENN.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008941; TPR-like.
DR InterPro; IPR005113; uDenn.
DR Pfam; PF03455; dDenn; 1.

```

DR Pfam; PF02141; DENN; 1.
DR Pfam; PF01535; PPR; 2.
DR Pfam; PF03456; UDENN; 1.
DR PROSITE; PS0947; DDENN; 1.
DR PROSITE; PS0211; DENN; 1.
DR PROSITE; PS0946; UDENN; 1.
FT NON_TER 928
SQ SEQUENCE 928 AA; 104964 MW; 7ACAD55D5204CB0B CRC64;

Query Match 40.4%; Score 57; DB 2; Length 928;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKPLSTASALL 20
: ||||| ||||| :|
Db 241 IECWPSNSKYPLPVSTFVL 260

RESULT 2
Q86T77 PRELIMINARY; PRT; 1831 AA.
AC Q86T77;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp451C1717 (Fragment).
GN Name=DKFZp451C1717;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human spinal cord;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Newes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL832602; CAD89960.1; -
DR InterPro; IPR005112; dDENN.
DR InterPro; IPR001194; DENN.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR005113; UDENN.
DR Pfam; PF03455; dDENN; 1.
DR Pfam; PF02141; DENN; 1.
DR Pfam; PF01535; PPR; 2.
DR Pfam; PF03456; UDENN; 1.
DR PROSITE; PS0947; DDENN; 1.
DR PROSITE; PS0211; DENN; 1.
DR PROSITE; PS0946; UDENN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1831 AA; 205650 MW; C9CD64C74C4A907A CRC64;

Query Match 40.4%; Score 57; DB 2; Length 1831;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKPLSTASALL 20
: ||||| ||||| :|
Db 209 IECWPSNSKYPLPVSTFVL 228

RESULT 3
Q81VX2 PRELIMINARY; PRT; 1863 AA.
AC Q81VX2;
DT 01-VAR-2003 (TrEMBLrel. 23, Created)
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-myc promoter binding protein.
GN Name=MYCBP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udwin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041706; AAH41706.1; -
DR InterPro; IPR005112; dDENN.
DR InterPro; IPR001194; DENN.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008941; TPR-like.
DR InterPro; IPR005113; UDENN.
DR Pfam; PF03455; dDENN; 1.
DR Pfam; PF02141; DENN; 1.
DR Pfam; PF01535; PPR; 2.
DR Pfam; PF03456; UDENN; 1.
DR PROSITE; PS0947; DDENN; 1.
DR PROSITE; PS0211; DENN; 1.
DR PROSITE; PS0946; UDENN; 1.
SQ SEQUENCE 1863 AA; 209226 MW; CE653AD6AA6749A1 CRC64;

Query Match 40.4%; Score 57; DB 2; Length 1863;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VSCWPSYLYKPLSTASALL 20
: ||||| ||||| :|
Db 241 IECWPSNSKYPLPVSTFVL 260

RESULT 4
Q7Z401 PRELIMINARY; PRT; 1865 AA.
AC Q7Z401;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-MYC promoter-binding protein IRLB.
GN Name=IRLB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22788788; PubMed=12906859;
RA Semova N., Kapanadze B., Corcoran M., Kutsenko A., Bazanova A.,
RA Semov A.;

```



```

RESULT 7
Q6GZ70
ID Q6GZ70 PRELIMINARY; PRT; 118 AA.
AC Q6GZ70;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Constitutive androstane receptor SV20 (Constitutive androstane
DE receptor SV19).
GN Name=NR1I3;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamba J.K., Lamba V., Yasuda K., Lin Y.S., Assem M., Thompson E.,
RA Strom S., Schuetz E.G.;
RT "Expression of CAR splice variants in human tissues and their
RT functional consequences.";
RL J. Pharmacol. Exp. Ther. 0:0-0(2004).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AY572825; AAT47178.1; -
DR EMBL; AY572824; AAT47177.1; -
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00105; zf-C4; 1.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00399; Znf_C4; 1.
DR DNA-binding; Metal-Binding; Nuclear protein; Receptor;
KW Transcription regulation; Zinc; Zinc-finger.
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 118 AA; 13460 MW; E279304F33688B46 CRC64;

Query Match 36.5%; Score 51.5; DB 2; Length 118;
Best Local Similarity 51.7%; Pred. No. 13;
Matches 15; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 4 WPSYLKX--PLS-----TASASLLATQLK 25
Db 67 WKSVTYSIPLSVSKHKTSAGLFATQLK 95

RESULT 8
Q6GZ72
ID Q6GZ72 PRELIMINARY; PRT; 147 AA.
AC Q6GZ72;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Constitutive androstane receptor SV18 (Constitutive androstane
DE receptor SV17).
GN Name=NR1I3;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamba J.K., Lamba V., Yasuda K., Lin Y.S., Assem M., Thompson E.,
RA Strom S., Schuetz E.G.;
RT "Expression of CAR splice variants in human tissues and their
RT functional consequences.";
RL J. Pharmacol. Exp. Ther. 0:0-0(2004).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AY572823; AAT47176.1; -
DR EMBL; AY572822; AAT47175.1; -
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PRO0047; STROIDFINGER.

```

```

DR PRINTS; PRO0350; VITAMINDR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00311; NUCLEAR RECEPTOR; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 147 AA; 16602 MW; F03F83F41D00D493 CRC64;

Query Match 36.5%; Score 51.5; DB 2; Length 147;
Best Local Similarity 51.7%; Pred. No. 17;
Matches 15; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 4 WPSYLKX--PLS-----TASASLLATQLK 25
Db 96 WKSVTYSIPLSVSKHKTSAGLFATQLK 124

RESULT 9
Q8ZJ53
ID Q8ZJ53 PRELIMINARY; PRT; 306 AA.
AC Q8ZJ53;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Type III secretion system apparatus protein.
GN OrderedLocusNames=YPO0269;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Brooks K., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -!- SIMILARITY: Belongs to the fln/mopA/spaO family.
DR EMBL; AJ414141; CAC89132.1; -
DR PIR; A10033; A10033.
DR GO; GO:0009425; C:flagellar basal body (sensu Bacteria); IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0005539; P:ciliary/flagellar motility; IEA.
DR InterPro; IPR001172; Flagellar_Flin.
DR InterPro; IPR001543; SpaA.
DR Pfam; PF01052; SpaA; 1.
DR PRINTS; PR00956; FLGMOTORFLIN.
DR ProDom; PD001777; SpaA; 1.
KW Complete proteome.
SQ SEQUENCE 306 AA; 33423 MW; 964683C5A79783A4 CRC64;

Query Match 36.2%; Score 51; DB 2; Length 306;
Best Local Similarity 43.5%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 SCWPSYLKYP LSTASASLLATQL 24
Db 41 SCQHWINTLATDNPQLLAEL 63

RESULT 10
Q8D1G2
ID Q8D1G2 PRELIMINARY; PRT; 320 AA.
AC Q8D1G2; Q74XJ4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

```

Query Match	36.2%	Score 51	DB 2	Length 320
Best Local Similarity	43.5%	Pred. No. 45		
Matches 10	Conservative	4	Mismatches	9
			Indels	0
			Gaps	0
QY	2	SCWPSYLKYPPLSTASASLLATQL 24		
DB	55	SCQHWLNTLTATDNPQLLAEL 77		
		: : : : :		
		: : : : :		
RESULT 12				
Q9JGK4				
ID	Q9JGK4	PRELIMINARY;	PRT;	394 AA.
AC	Q9JGK4;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Long Req.			
GN	Name=meq;			
OS	Meleagrid herpesvirus 1 (MeHV-1) (Turkey herpesvirus)			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=10390;			
	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=CVI988;			
RC	Lee S., Takagi M., Ohashi K., Sugimoto C., Onuma M.;			
RRL	Submitted (OCT-1999) to the EMBL/GenBank/DBSJ databases.			
RL	-!- SIMILARITY: Belongs to the BZIP family.			
CCC	EMBL; AB033119; BAA93716.1; ..			
DR	GO; GO:0005634; C:nucleus; IEA.			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0003355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR008917; Euk.transc._DNA.			
DR	InterPro; IPR000837; LeuZip_Pos.			
DR	InterPro; IPR009053; Prefoldin.			
DR	InterPro; IPR004827; TF_BZIP.			
DR	Pfam; PF001070; BZIP; 1.			
DR	PRINTS; PRO0042; LEUZIPPRFOS.			
DR	SMART; SM00338; BRLZ; 1.			
DR	PROSITE; PS0217; BZIP; 1.			
DR	PROSITE; PS00036; BZIP_BASIC; 1.			
KW	DNA-binding; Nuclear protein.			
SW	SEQUENCE 394 AA; 42830 MW; FF95DDEE786450F6 CRC64;			
QY	1	VSCWPSYLKYPPLSTASAS 18		
DB	337	LSCWPGSPFYPISTSGS 354		
		: : : : :		
		: : : : :		
RESULT 13				
P73143				
ID	P73143	PRELIMINARY;	PRT;	565 AA.
AC	P73143;			
DT	01-FEB-1997 (TrEMBLrel. 02, Created)			
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Slr1033 protein.			
GN	OrderedLocusNames=slr1033;			
OS	Synechocystis sp. (strain PCC 6803).			
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.			
OX	NCBI_TaxID=1148;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96127529; PubMed=8590279;			
RA	Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,			
RA	Sugita M., Tabata S.;			
RT	"sequence analysis of the genome of the unicellular cyanobacterium			
RT	Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb			

```

RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirasawa M., Sugura M., Sasamoto S., Kimura T.,
RA Hosouchi I., Takeuchi A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996).
DR ZMBL; D90904; BAAL17183.1; -.
DR FIR; S75255; S75255.
KW Complete proteome.
SQ SEQUENCE 565 AA; 64181 MW; 1C569961D4ACBF90 CRC64;

Query Match 35.5%; Score 50; DB 2; Length 565;
Best Local Similarity 36.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 4 WPSYLYKPLSTASASLLATQLKSLA 28
DB 316 WTQVLNPLTPREANILATEIASLS 340

RESULT 14
Q7X329 PRELIMINARY; PRT; 608 AA.
AC Q7X329;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS uncultured Acidobacteria bacterium.
OC Bacteria; Acidobacteria; environmental samples.
OX NCBI_TaxID=171953;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22983092; PubMed=14617179;
RA Quaiser A., Ochsenreiter T., Ianz C., Schuster S.C., Treusch A.H.,
RA Eck J., Schleper C.;
RT "Acidobacteria form a coherent but highly diverse group within the
RT bacterial domain: evidence from environmental genomics.";
RL Mol. Microbiol. 50:563-575(2003).
DR ZMBL; AY281354; AAP58535.1; -.
KW Hypothetical protein.
SQ SEQUENCE 608 AA; 64406 MW; 2C98C98E9BF21AA9 CRC64;

Query Match 35.5%; Score 50; DB 2; Length 608;
Best Local Similarity 52.4%; Pred. No. 1.3e+02;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 WPSYLYKPLSTASASLLATQL 24
DB 37 WVTLLSFPLSAAAYLLARHL 57

RESULT 15
O47500 PRELIMINARY; PRT; 760 AA.
AC O47500;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE RT-like protein.
GN Name=cytb;
OS Venturia inaequalis (Apple scab fungus).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;

```

```

OC Pleosporales; Venturiaceae; Venturia.
OX NCBI_TaxID=5025;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S-56-88;
RX MEDLINE=98041928; PubMed=9371888;
RA Zheng D., Koller W.;
RT "Characterization of the mitochondrial cytochrome b gene from Venturia
RT inaequalis.";
RL Curr. Genet. 32:361-366(1997).
DR ZMBL; AF004559; AAB95256.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008380; P:RNA splicing; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005797; Cytb_b6 N.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR000442; Intron_maturase2.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN 1.
KW Mitochondrion; RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 760 AA; 85883 MW; 61B0C1D5E6C55EA2 CRC64;

Query Match 35.5%; Score 50; DB 2; Length 760;
Best Local Similarity 55.0%; Pred. No. 1.6e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 6 SYLYKPLSTASASLLATQLK 25
DB 678 SYLEYILKQSCAKLLATPK 697

Search completed: November 10, 2004, 15:53:35
Job time : 49.0093 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14, Search time 8.78491 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-51

Perfect score: 138

Sequence: 1 VHNFRHWGLPLSLNLYPLFLSP 24

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.5	38.0	3165	2	US-08-459-146-3
2	52.5	38.0	3165	2	US-08-459-065-3
3	50	36.2	470	4	US-09-134-000C-4816
4	49	35.5	531	4	US-09-489-039A-9781
5	48	34.8	287	3	US-09-134-001C-4332
6	48	34.8	287	4	US-09-710-279-2320
7	48	34.8	420	4	US-09-328-352-5907
8	48	34.8	435	4	US-09-489-039A-13740
9	48	34.8	785	4	US-09-079-030-216
10	47	34.1	114	4	US-09-248-796A-16735
11	47	34.1	273	4	US-09-252-991A-29785
12	47	34.1	335	4	US-09-328-352-5349
13	47	34.1	920	1	US-08-451-715A-2
14	46	33.3	77	4	US-09-134-000C-6358
15	46	33.3	100	4	US-09-079-030-43
16	46	33.3	215	3	US-09-134-001C-5483
17	46	33.3	278	4	US-09-710-279-2912
18	46	33.3	394	4	US-08-079-030-833
19	46	33.3	773	4	US-09-079-030-215
20	46	33.3	840	4	US-09-079-030-214
21	46	33.3	4536	4	US-08-180-422B-27
22	46	33.3	4536	4	US-09-079-030-1
23	46	33.3	4563	4	US-09-108-006C-1
24	46	33.3	4563	4	US-09-538-092-842
25	45	32.6	150	4	US-08-198-452A-989
26	45	32.6	294	2	US-08-923-856-3
27	45	32.6	294	3	US-09-216-294-3

28	45	32.6	370	1	US-08-415-818-12	Sequence 12, Appl
29	45	32.6	370	2	US-08-894-236-12	Sequence 12, Appl
30	45	32.6	370	5	PCT-US96-01444-12	Sequence 12, Appl
31	45	32.6	394	4	US-09-270-767-41366	Sequence 41366, A
32	45	32.6	394	4	US-09-270-767-56582	Sequence 56582, A
33	45	32.6	533	4	US-09-252-991A-21407	Sequence 21407, A
34	45	32.6	1169	1	US-08-315-468-4	Sequence 4, Appl
35	44.5	32.2	297	3	US-09-173-581-1	Sequence 1, Appl
36	44.5	32.2	297	3	US-09-420-915-1	Sequence 1, Appl
37	44.5	32.2	628	4	US-09-107-532A-5288	Sequence 5288, Ap
38	44.5	32.2	1441	4	US-09-252-991A-28143	Sequence 28143, A
39	44	31.9	168	4	US-09-710-279-724	Sequence 724, App
40	44	31.9	170	3	US-09-134-001C-3512	Sequence 3512, Ap
41	44	31.9	226	4	US-09-869-388-10	Sequence 10, Appl
42	44	31.9	227	4	US-09-869-388-6	Sequence 6, Appl
43	44	31.9	381	4	US-09-489-039A-12111	Sequence 12111, A
44	44	31.9	450	4	US-09-765-111A-25	Sequence 25, Appl
45	44	31.9	464	4	US-09-252-991A-25514	Sequence 25514, A

ALIGNMENTS

RESULT 1
US-08-459-146-3
; Sequence 3, Application US/08459146
; Patent No. 5866405
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,146
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptophnetria
; ORGANISM: parasitica)
; STRAIN: EP713
; US-08-459-146-3

Query Match 38.0%; Score 52.5; DB 2; Length 3165;

Best Local Similarity 54.2%; Pred. No. 86;
Matches 13; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 VHNFGHWGLPLSLFLNYPLFLSP 24
Db 514 VRNTGIH-GLPGDFLSNYPRLPTP 536

RESULT 2

US-08-459-065-3
; Sequence 3, Application US/08459065
; Patent No. 5882642
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,065
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3165 amino acids
; TYPE: amino acid

TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptophnectria
; ORGANISM: parasitica)
; STRAIN: EP713
US-08-459-065-3

Query Match 38.0%; Score 52.5; DB 2; Length 3165;
Best Local Similarity 54.2%; Pred. No. 86;
Matches 13; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 VHNFGHWGLPLSLFLNYPLFLSP 24
Db 514 VRNTGIH-GLPGDFLSNYPRLPTP 536

RESULT 3

US-09-134-000C-4816
; Sequence 4816, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4816
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4816

Query Match 36.2%; Score 50; DB 4; Length 470;
Best Local Similarity 40.9%; Pred. No. 25;
Matches 9; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 4 FGRHWGLPLS--FLNYPFLS 23
Db 88 YGHRWGARYSWYWINYPLMLA 109

RESULT 4

US-09-489-039A-9781
; Sequence 9781, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9781
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9781

Query Match 35.5%; Score 49; DB 4; Length 531;
Best Local Similarity 56.7%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFL 16
Db 303 GNEWGLPFNPLL 314

RESULT 5

US-09-134-001C-4332
; Sequence 4332, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4332
; LENGTH: 287
; TYPE: PRT

US-09-134-001C-4332
; ORGANISM: Staphylococcus epidermidis

```

; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13740
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13740

Query Match 34.8%; Score 48; DB 4; Length 435;
Best Local Similarity 32.5%; Pred.No. 44;
Matches 13; Conservative 3; Mismatches 6; Indels 18; Gaps 2;

Qy 3 NFGRRH---GLPLSFL-----LNYPFLFLSP 24
Db 249 NFLHWRDRGLPLFVEGFLMGAFVTLFNYGIRLMMSP 288

RESULT 9
US-09-079-030-216
; Sequence 216, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogeveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; TITLE OF INVENTION: VECTORS FOR TRANSECTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ARAG:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 785 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-079-030-216

Query Match 34.8%; Score 48; DB 4; Length 785;
Best Local Similarity 54.5%; Pred.No. 85;
Matches 12; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

Qy 9 GLPL-----SFLNYPFLFLSP 24

```


US-09-134-000C-6358
; Sequence 6358, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6358
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6358

Query Match 33.3%; Score 46; DB 4; Length 77;
Best Local Similarity 90.0%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 LSFLLNYPLF 21
| | | | | | | |
Db 51 LFLNYPLF 60

RESULT 15
US-09-079-030-43
; Sequence 43, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogeveen, Ron C.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ABAG:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-079-030-43

Query Match 33.3%; Score 46; DB 4; Length 100;
Best Local Similarity 69.2%; Pred. No. 18;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 12 LSFLLNYPLF 24
: | | | | | | | |
Db 31 IDFLNNYALFLSP 43

Search completed: November 10, 2004, 14:55:37
Job time : 8.83491 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match Length	%			
1	138	100.0	24	14	US-10-092-750-51	Sequence 51, Appl
2	60	43.5	1447	9	US-09-797-097-2	Sequence 2, Appli
3	54.5	39.5	374	15	US-10-424-599-183840	Sequence 2, Appli
4	54	39.1	88	16	US-10-437-963-104238	Sequence 183840,
5	54	39.1	461	15	US-10-282-122A-60769	Sequence 104238,
6	54	39.1	571	14	US-10-369-736-3	Sequence 60769, A
7	54	39.1	571	14	US-10-369-736-49	Sequence 3, Appli
8	54	39.1	571	14	US-10-369-738-3	Sequence 49, Appl
9	54	39.1	571	14	US-10-369-738-49	Sequence 3, Appli
10	51.5	37.3	372	15	US-10-282-122A-78114	Sequence 49, Appl
11	51	37.0	61	15	US-10-424-599-216105	Sequence 78114, A
12	51	37.0	284	16	US-10-437-963-185786	Sequence 216105,
13	50	36.2	458	15	US-10-282-122A-57060	Sequence 185786,
						Sequence 57060, A

us-10-092-750-51.rapb

Fri Nov 12 14:55:36 2004

```

; APPLICANT: All, Benjamin
; APPLICANT: Howard, John
; TITLE OF INVENTION: Expression Cassettes and Methods of Delivery of Animal Vaccines
; FILE REFERENCE: P00246US1
; CURRENT APPLICATION NUMBER: US/09/797,097
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/529,006
; PRIOR FILING DATE: 1995-09-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Transmissible Gastroenteritis Virus Spike protein
US-09-797-097-2

Query Match      43.5%; Score 60; DB 9; Length 1447;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHNLLETFLNYSRLSP 49

RESULT 3
US-10-424-599-183840
; Sequence 183840, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 183840
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137021C.1.pep
US-10-424-599-183840

Query Match      39.5%; Score 54.5; DB 15; Length 374;
Best Local Similarity 52.4%; Pred. No. 19;
Matches 11; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 3 NFGRHGWLPLSFLNYPFLS 23
DB 308 NYGTHGQQV-FLLHPFMLS 327

RESULT 4
US-10-437-963-104238
; Sequence 104238, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

```

```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104238
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(88)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101591C.1.pep
US-10-437-963-104238

Query Match      39.1%; Score 54; DB 16; Length 88;
Best Local Similarity 56.5%; Pred. No. 5.1;
Matches 13; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 8 WGLPLSFLN-----YPLFLSP 24
DB 2 WLPLSTLLSTSLFLYPLFLSP 24

RESULT 5
US-10-282-122A-60769
; Sequence 60769, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60769
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60769

```


QY 2 HNFCGRHWGLPLSLFLLNYPL 20
| | | | | | | | | | | |
7 HPEWETUNCIILIIILEAPL 25

QY 2 HNFGRHWGLPLSFLLNYPL 20
| | | | | | | | | |
2 VDFRHWGLPLSFLLNYPL 25

```

RESULT 10
US-10-382-122A-78114
; Sequence 78114, Application US/10282122A
; Publication No. US2004009129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

```


Search completed: November 11, 2004, 02:43:07
Job time : 28.8085 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.88679 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-51
Perfect score: 138
Sequence: 1 VHNFRHWGLPLSLNYPLFLSP 24
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	43.5	1449	1 A43573	E2 glycoprotein pr
2	60	43.5	1449	1 VGIHFS	E2 glycoprotein pr
3	55	39.9	1449	2 G7423	E2 glycoprotein pr
4	54	38.1	461	2 AF1079	amino acid transpo
5	52.5	38.0	3165	2 S15010	hypothetical prote
6	51.5	37.3	372	2 AB0276	probable membrane
7	51	37.0	217	2 H98155	probable permealase
8	51	37.0	217	2 AC3132	hypothetical prote
9	51	37.0	467	2 AC2015	hypothetical prote
10	49.5	35.9	796	2 T15424	hypothetical prote
11	49	35.5	494	2 AC0133	probable permealase
12	48.5	35.1	821	2 B70199	outer membrane pro
13	48	34.8	417	2 AB0682	probable membrane
14	48	34.8	417	2 F90916	probable transport
15	48	34.8	417	2 F64915	membrane protein y
16	48	34.8	417	2 C85765	probable transport
17	48	34.8	448	2 CF1942	hypothetical prote
18	48	34.8	784	2 JH0101	apolipoprotein B-1
19	47	34.1	134	2 B46178	probable transcrip
20	47	34.1	451	2 S66008	transport protein
21	47	34.1	477	2 AC2834	UDP-MurNAc-pentape
22	47	34.1	477	2 G97611	UDP-N-acetylmutamo
23	47	34.1	650	2 T23175	hypothetical prote
24	47	34.1	920	2 F64697	isoleucine-tRNA li
25	47	34.1	920	2 F71823	isoleucine-tRNA li
26	47	34.1	1398	2 T21884	hypothetical prote
27	47	34.1	1447	1 VGIHFS	E2 glycoprotein pr
28	47	34.1	1447	1 VGIHFS	E2 glycoprotein pr
29	46.5	33.7	252	2 H90191	hypothetical prote

hypothetical prote
hypothetical prote
conserved hypothet
D-alanine-D-alanin
related to UDP N-A
probable exporter
related to trfA pr
brefeldin A-sensit
apolipoprotein B -
apolipoprotein B-1
hypothetical prote
hypothetical prote
hypothetical prote
UDP-N-acetylmuramo
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A43573
E2 glycoprotein precursor - porcine transmissible gastroenteritis virus (strain Miller)
N;Alternate names: Peplomer glycoprotein; spike glycoprotein
C;Species: porcine transmissible gastroenteritis virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A43573
R;Wesley, R.D.
Adv. Exp. Med. Biol. 276, 301-306, 1990
A;Title: Nucleotide sequence of the E2-peplomer protein gene and partial nucleotide se
A;Reference number: A43573; MUID:91353366; PMID:1966416
A;Accession: A43573
A;Molecule type: genomic RNA
A;Residues: 1-1449 <RES>
A;Cross-references: UNIPROT:P33470; GB:S51223; NID:G234109; PIDN:AAB19567.1; PID:G23411
A;Note: the authors translated the codon GAA for residue 388 as Cys
C;Superfamily: coronavirus E2 glycoprotein
C;Keywords: glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1449/Product: E2 glycoprotein #status predicted <E2G>
F;1027-1043/Region: hydrophobic
F;1391-1411/Domain: transmembrane #status predicted <TMN>
F;26,42,71,94,243,250,285,334,345,362,375,405,449,516,532,554,594,704,725,780,819,834,
d

Query Match 43.5%; Score 60; DB 1; Length 1449;
Best Local Similarity 60.0%; Pred. No. 1.3;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 GRHWGLPLSLNYPLFLSP 24
| | | | | : | | | | |
Db 30 GNEWNLTETPLNYSRLSP 49

RESULT 2

VGIHFS
E2 glycoprotein precursor - porcine transmissible gastroenteritis virus (strain FS772//
N;Alternate names: Peplomer glycoprotein; spike glycoprotein
C;Species: porcine transmissible gastroenteritis virus
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: B43489; S11728
R;Britton, P.; Page, K.W.
Virus Res. 18, 71-80, 1990
A;Title: Sequence of the S gene from a virulent British field isolate of transmissible
A;Reference number: A43489; MUID:91188698; PMID:1964522
A;Accession: B43489
A;Molecule type: mRNA
A;Residues: 1-1449 <BRI>
A;Cross-references: UNIPROT:P18450; GB:X53128; NID:G61377; PIDN:CAA37285.1; PID:G61379
C;Superfamily: coronavirus E2 glycoprotein
C;Keywords: Glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>

F:17-1449/Product: E2 glycoprotein #status predicted <E2G>
 F:1027-1043/Region: hydrophobic
 F:1395-1411/Domain: transmembrane #status predicted <TMN>
 F:26,42,71,94,243,250,285,334,345,362,375,405,449,516,532,554,594,704,725,780,819,834,84
 d

Query Match 43.5%; Score 60; DB 1; Length 1449;
 Best Local Similarity 60.0%; Pred. No. 1.3;
 Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPLFLSP 24
 DB 30 GNHWNLIETFLNYSRLSP 49

RESULT 3

S47423
 E2 glycoprotein precursor - porcine transmissible gastroenteritis virus

N;Alternate names: envelope protein; spike protein
 C;Species: porcine transmissible gastroenteritis virus
 C;Date: 23-Nov-1994 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
 C;Accession: S65851; S47423
 R;Chen, C.M.; Cavanagh, D.; Britton, P.
 Virus Res. 38, 83-89, 1995
 A;Title: Cloning and sequencing of a 8.4-kb region from the 3'-end of a Taiwanese virule
 A;Reference number: S65850; MUID:96060227; PMID:8546012
 A;Accession: S65851
 A;Status: preliminary
 A;Molecule type: nucleic acid sequence not shown; translation not shown
 A;Residues: 1-1449 <CH2>
 A;Cross-references: UNIPROT:Q88510; EMBL:Z35758; NID:G529246; PIDN:CAA84806.1; PID:G5292
 A;Experimental source: Taiwanese field isolate
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
 C;Genetics:
 A;Gene: S

C;Superfamily: coronavirus E2 glycoprotein
 C;Keywords: glycoprotein; transmembrane protein
 F:1-16/Domain: signal sequence #status predicted <STG>
 F:17-1449/Product: E2 glycoprotein #status predicted <E2G>

Query Match 39.9%; Score 55; DB 2; Length 1449;
 Best Local Similarity 55.0%; Pred. No. 6.9;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPLFLSP 24
 DB 30 GNHWNLIETFLNYSRLSP 49

RESULT 4

AF1079
 amino acid transporter homolog lmo0037 [imported] - Listeria monocytogenes (strain EGD-e

C;Species: Listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AF1079
 R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 ; Dominguez-Bernal, G.; Duchaud, E.; Duran, L.; Dussurget, O.; Enrian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schluetter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AF1079
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-461 <GLA>

A;Cross-references: UNIPROT:Q8YAS6; GB:NC_003210; PIDN:CAC98252.1; PID:gl6409396; GSPDB:
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: lmo0037

Query Match 39.1%; Score 54; DB 2; Length 461;

Best Local Similarity 40.9%; Pred. No. 3;
 Matches 9; Conservative 8; Mismatches 3; Indels 2; Gaps 1;
 QY 4 FGRHWGLPLSFL--LNYPFLS 23
 DB 76 FGRKGGARVAVLYWYINPIWMA 97

RESULT 5

SI5010

hypothetical protein B - Cryphonectria hypovirus 1

C;Species: Cryphonectria hypovirus 1
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
 C;Accession: SI5010
 R;Shapira, R.; Choi, G.H.; Nuss, D.L.
 EMBO J. 10, 731-739, 1991
 A;Title: Virus-like genetic organization and expression strategy for a double-stranded
 A;Reference number: SI5009; MUID:91184117; PMID:2009854
 A;Accession: SI5010
 A;Status: preliminary
 A;Molecule type: genomic RNA
 A;Residues: 1-3165 <EMB>
 A;Cross-references: UNIPROT:Q04350; GB:M57938; NID:G331157; PIDN:AAA67458.1; PID:G33115

Query Match 38.0%; Score 52.5; DB 2; Length 3165;
 Best Local Similarity 54.2%; Pred. No. 37;
 Matches 13; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 VHNFGRWGLPLSFLNYPLFLSP 24
 DB 514 VRNTGIH-GUGDFLSNYPLRPTP 536

RESULT 6

AB0276

probable membrane protein YPO2266 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AB0276
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AB0276
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-372 <KUR>

A;Cross-references: UNIPROT:Q8ZEB5; GB:AL590842; PIDN:CAC91070.1; PID:gl5980262; GSPDB:
 C;Genetics:
 C;Superfamily: probable antibiotic resistance protein YybF

Query Match 37.3%; Score 51.5; DB 2; Length 372;
 Best Local Similarity 35.6%; Pred. No. 5.6;
 Matches 16; Conservative 1; Mismatches 5; Indels 23; Gaps 3;

QY 3 NFGRW--GLPLSF-----LLNY-----PLFLSP 24
 DB 242 NFKLWRDGLPLFLLAAGFLMGVVTFLNYIGYLLDGFYLLSP 286

RESULT 7

H98155

Probable permealase of ABC transporter PAL1257 [imported] - Agrobacterium tumefaciens (str

C;Species: Agrobacterium tumefaciens
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C;Accession: H98155
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu

A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H98155
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <KUR>
A;Cross-references: UNIPROT:Q8UXK6; GB:AE007870; PIDN:AAK88770.1; PID:g15158516; GSPDB:C
C;Genetics:
A;Gene: AGR_L403
A;Map position: linear chromosome
C;Superfamily: histidine permease protein M

Query Match 37.0%; Score 51; DB 2; Length 217;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 FGRHWGLPLSLN 17
||:|||||:
DB 46 FGPMWLSLPLTVLD 59

RESULT 8

AD3132
Hypothetical protein Atu4680 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A;Accession: AD3132
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD3132
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <KUR>
A;Cross-references: UNIPROT:Q8UXK6; GB:AE008689; PIDN:AAL45474.1; PID:g17743180; GSPDB:C
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4680
A;Map position: linear chromosome
C;Superfamily: histidine permease protein M

Query Match 37.0%; Score 51; DB 2; Length 217;
Best Local Similarity 57.1%; Pred. No. 3.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 FGRHWGLPLSLN 17
||:|||||:
DB 46 FGPMWLSLPLTVLD 59

RESULT 9

AC2015
Hypothetical protein all1673 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
A;Accession: AC2015
R;Kareko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <KUR>
A;Cross-references: UNIPROT:Q8YNE0; GB:BA000019; PIDN:BAW8039.1; PID:g17135493; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:

A;Gene: all1673

Query Match 37.0%; Score 51; DB 2; Length 467;
Best Local Similarity 32.1%; Pred. No. 8.4;
Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;

QY 1 VHNFGRHGWLPLS-----FLNLYPLFL 22
||:|||||:
DB 412 LEGLSRHGWGLFANRDFIFLFIATFPVIL 439

RESULT 10

T16424
Hypothetical protein F52C9.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
C;Accession: T16424
R;Favell, T.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F52C9.
A;Reference number: Z18511
A;Accession: T16424
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-796 <FAV>
A;Cross-references: EMBL:U39850; NID:g1055052; PID:g1055060; PIDN:AAA81062.1; CESP:F52C
C;Genetics:
A;Gene: CESP:F52C9.1
A;Introns: 98/1; 265/3; 354/3; 407/3; 524/2; 565/2; 586/1; 721/3; 726/1; 748/2
C;Superfamily: Caenorhabditis elegans hypothetical protein F52C9.1

Query Match 35.9%; Score 49.5; DB 2; Length 796;
Best Local Similarity 44.0%; Pred. No. 25;
Matches 11; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 VHNFGRHGWLPLSFLNLYPLFLSP 24
||:|||||:
DB 421 VDRFAKMGVPLSSFCDDFAELIP 445

RESULT 11

AC0133
Probable permease YPO1083 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0133
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.E
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-494 <KUR>
A;Cross-references: UNIPROT:Q8ZH28; GB:AL590842; PIDN:CAC89926.1; PID:g15979151; GSPDB:
C;Genetics:
A;Gene: YPO1083

Query Match 35.5%; Score 49; DB 2; Length 494;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLN 16
||:|||||:
DB 267 GNRWGLFNFLL 278

RESULT 12

B70199
outer membrane protein homolog - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: B70199
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervilavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: B70199
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-821 <KLE>
A;Cross-references: UNIPROT:O51735; GB:AE001178; GB:AE000783; NID:g2688731; PIDN:AAC6713
A;Experimental source: strain B31

Query Match 35.1%; Score 48.5; DB 2; Length 821;
Best Local Similarity 35.7%; Pred. No. 36;
Matches 10; Conservative 4; Mismatches 5; Indels 9; Gaps 1;

QY 3 NFGGRHWG-----LPLSFLNLYPLF 21
DB 768 SFATSGTGIRLLPQLPLSVIAIPFY 795

RESULT 13
AB0682
Probable membrane transport protein STY1579 [imported] - *Salmonella enterica* subsp. ente
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: this species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AB0682
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB0682
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-417 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01827.1; PID:g16502671; GSPDB:GN00176
C;Genetics:
A;Gene: STY1579
C;Superfamily: probable antibiotic resistance protein yybF

Query Match 34.8%; Score 48; DB 2; Length 417;
Best Local Similarity 35.0%; Pred. No. 21;
Matches 14; Conservative 1; Mismatches 7; Indels 18; Gaps 2;

QY 3 NFGGRHWG-----GLPLSF-----LLNYPFLFLSP 24
DB 230 NFRLLWRDRGLPLFAEGFLMGAFVTLFNYIGYRLMLSP 269

RESULT 14
F90916
Probable transport protein ECs2302 [imported] - *Escherichia coli* (strain O157:H7, subst
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90916
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90916
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-417 <HAY>
A;Cross-references: UNIPROT:Q8X784; GB:BA000007; PIDN:BA035725.1; PID:g13361769; GSPDB:G

A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2302
C;Superfamily: probable antibiotic resistance protein yybF

Query Match 34.8%; Score 48; DB 2; Length 417;
Best Local Similarity 35.0%; Pred. No. 21;
Matches 14; Conservative 1; Mismatches 7; Indels 18; Gaps 2;

QY 3 NFGGRHWG-----GLPLSF-----LLNYPFLFLSP 24
DB 230 NFRLLWRDRGLPLFAEGFLMGFSVTLFNYIGYRLMLSP 269

RESULT 15
F64915
membrane protein ynfM - *Escherichia coli* (strain K-12)
C;Species: *Escherichia coli*
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: F64915
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A;Rose, D.J.; May, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of *Escherichia coli* K-12.
A;Reference number: A64720; MUID:97425617; PMID:9278503
A;Accession: F64915
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-417 <BLAT>
A;Cross-references: UNIPROT:P43531; GB:AE000255; GB:U00096; NID:g1787875; PIDN:AAC74668
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ynfM
C;Superfamily: probable antibiotic resistance protein yybF
C;Keywords: inner membrane; transmembrane protein; transport protein
F;35-55/Domain: transmembrane #status predicted <TM1>
F;75-91/Domain: transmembrane #status predicted <TM2>
F;103-119/Domain: transmembrane #status predicted <TM3>
F;127-143/Domain: transmembrane #status predicted <TM4>
F;191-207/Domain: transmembrane #status predicted <TM5>
F;242-258/Domain: transmembrane #status predicted <TM6>
F;313-329/Domain: transmembrane #status predicted <TM7>
F;392-408/Domain: transmembrane #status predicted <TM8>

Query Match 34.8%; Score 48; DB 2; Length 417;
Best Local Similarity 35.0%; Pred. No. 21;
Matches 14; Conservative 1; Mismatches 7; Indels 18; Gaps 2;

QY 3 NFGGRHWG-----GLPLSF-----LLNYPFLFLSP 24
DB 230 NFRLLWRDRGLPLFAEGFLMGFSVTLFNYIGYRLMLSP 269

Search completed: November 10, 2004, 14:52:20
Job time : 6.88679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 31.9698 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-51

Perfect score: 138

Sequence: 1 VHNFGHWGLPLSLNYPFLSP 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	45.7	475	2 Q7QJ82	Q7QJ82 anopheles g
2	60	43.5	269	2 Q6QIL4	Q6QIL4 transmissib
3	60	43.5	269	2 AAS68178	AAS68178 transmiss
4	60	43.5	1449	1 VGL2_CVPFS	P18450 porcine tra
5	60	43.5	1449	1 VGL2_CVPFI	P33470 porcine tra
6	60	43.5	1449	2 G6PQ05	G6PQ05 transmissib
7	60	43.5	1449	2 Q7T428	Q7T428 transmissib
8	60	43.5	1449	2 AAT00645	AAT00645 transmiss
9	59	42.8	808	2 Q9QRQ0	Q9QRQ0 canine coro
10	55	39.9	184	2 Q41204	Q41204 transmissib
11	55	39.9	789	2 Q8QRR7	Q8QRR7 transmissib
12	55	39.9	789	2 Q8QRR9	Q8QRR9 transmissib
13	55	39.9	789	2 Q8QRR0	Q8QRR0 transmissib
14	55	39.9	789	2 Q8QRR3	Q8QRR3 transmissib
15	55	39.9	790	2 Q8QRR2	Q8QRR2 transmissib
16	55	39.9	1449	2 Q84853	Q84853 porcine res
17	55	39.9	1449	2 Q88510	Q88510 transmissib
18	55	39.9	1449	2 Q9DY22	Q9DY22 transmissib
19	54	39.1	59	2 Q8K312	Q8K312 mus musculu
20	54	39.1	207	2 Q6QID2	Q6QID2 coltivirus
21	54	39.1	207	2 Q6QID3	Q6QID3 coltivirus
22	54	39.1	207	2 Q9YWN8	Q9YWN8 banana viru
23	54	39.1	207	2 AAS68264	AAS68264 coltivirus
24	54	39.1	207	2 AAS68265	AAS68265 coltivirus
25	54	39.1	260	2 Q85085	Q85085 transmissib
26	54	39.1	299	2 Q8QEW3	Q8QEW3 transmissib
27	54	39.1	461	2 Q8YAS6	Q8YAS6 listeria mo
28	54	39.1	461	2 Q725C7	Q725C7 listeria mo
29	54	39.1	461	2 AAT02834	AAT02834 listeria
30	54	39.1	467	2 Q6WRY0	Q6WRY0 xenopus tro
31	54	39.1	467	2 AAQ02688	AAQ02688 xenopus t

32 54 39.1 571 2 Q7TQN3 Q7TQN3 mus musculu
33 52.5 38.0 3164 2 Q9YTU2 Q9YTU2 cryphonectr
34 52.5 38.0 3165 2 Q04350 Q04350 cryphonectr
35 51.5 37.3 372 2 Q8D0I2 Q8D0I2 versinia pe
36 51.5 37.3 372 2 Q8ZEB5 Q8ZEB5 versinia pe
37 51 37.0 164 2 Q88WE9 Q88WE9 lactobacilli
38 51 37.0 217 2 Q7CVK4 Q7CVK4 agrobacteri
39 51 37.0 217 2 Q8U6X6 Q8U6X6 agrobacteri
40 51 37.0 347 2 Q8XSF7 Q8XSF7 ralstonia s
41 51 37.0 467 1 PCXA_ANASP Q8YWE0 anabaena sp
42 51 37.0 1449 2 Q9YRA7 Q9YRA7 transmissib
43 50.5 36.6 1223 2 Q7KWX7 Q7KWX7 dictyosteli
44 50.5 36.6 1223 2 AAS38666 AAS38666 dictyoste
45 50 36.2 319 2 Q8VGY6 Q8VGY6 mus musculu

ALIGNMENTS

RESULT 1

Q7QJ82

ID Q7QJ82 PRELIMINARY; PRT; 475 AA.

AC Q7QJ82;

DT 01-MAR-2004 (TrEMBLrel. 26, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ACP3645 (Fragment)

GN Name=agCG57535; ORFNames=ENRANGG00000019047;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.

OC NCBI_TaxID=180454;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PEST;

RA Anopheles Genome Sequencing Consortium;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

-!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AAB0100807; EAA04057.1; --

DR GO; GO:0016020; Cmembrane; IEA.

DR GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.

DR GO; GO:0006865; P:amino acid transport; IEA.

DR InterPro; IPR002422; AA/rel_permease2.

DR Pfam; PF01490; Aa_trans; 1.

FT NON_TER 1

SQ SEQUENCE 475 AA; 52632 MW; BE0A5AD7F2A9C025 CRC64;

Query Match 45.7%; Score 63; DB 2; Length 475;

Best Local Similarity 45.0%; Pred. No. 1.5;

Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 4 FGRHWGLPLSLNYPFLS 23

Db 151 WGRWGMPTSYCIRYGLFIT 170

RESULT 2

Q6QIL4

ID Q6QIL4 PRELIMINARY; PRT; 269 AA.

AC Q6QIL4;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Spike glycoprotein (Fragment).

OS Transmissible gastroenteritis virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Coronavirus; Group 1 species.

OC NCBI_TaxID=11149;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HN2002;

```

RA He K., Ni Y.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RR EMBL; AY539716; AAS68178.1; -.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 30273 MW; 5CBECDD997759501 CRC64;

Query Match 43.5%; Score 60; DB 2; Length 269;
Best Local Similarity 60.0%; Pred. No. 2.4;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
Db 13 GHNWLIETFLNYSRLSP 32

RESULT 3
AAS68178 PRELIMINARY; PRT; 269 AA.
AC AAS68178;
DT 29-MAR-2004 (TREMELrel. 27, Created)
DT 29-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 29-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Spike glycoprotein (Fragment).
OC Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN2002;
RA He K., Ni Y.;
RT "Transmissible gastroenteritis virus strain HN2002 spike glycoprotein
RT (S) gene."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RR EMBL; AY539716; AAS68178.1; -.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 30273 MW; 5CBECDD997759501 CRC64;

Query Match 43.5%; Score 60; DB 2; Length 269;
Best Local Similarity 60.0%; Pred. No. 2.4;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
Db 13 GHNWLIETFLNYSRLSP 32

RESULT 4
VGL2_CVPFS STANDARD; PRT; 1449 AA.
AC P18450; Q85088;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (peplomer protein).
GN Name=S;
OS Porcine transmissible gastroenteritis coronavirus (strain FS772/70)
OS (TGEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11150;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9118698; PubMed=1964522;
RA Britton P., Page K.W.;
RT "Sequence of the S gene from a virulent British field isolate of
RT transmissible gastroenteritis virus.";
RL Virus Res. 18:71-80(1990).
CC -!- FUNCTION: The peplomer protein mediates the binding of virions to
CC the host cell receptor and is involved in membrane fusion and in
CC syncytium formation.

```

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; X53128; CAA37285.1; -.
 DR FIR; B43489; VGIHFS.
 DR InterPro; IPR002551; Corona_S1.
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01600; Corona_S1; 1.
 DR Pfam; PF01601; Corona_S2; 1.
 KW Envelope protein; Glycoprotein; Signal; Transmembrane.

FT SIGNAL 1 16
 FT CHAIN 17 1449 E2 glycoprotein.
 FT DOMAIN 17 1390 Extracellular (Potential).
 FT TRANSMEM 1391 1410 Potential.
 FT DOMAIN 1411 1449 Cytoplasmic (Potential).
 FT DOMAIN 1411 1432 Cys-rich.
 FT CARBOHYD 26 26 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 71 71 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 94 94 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 243 243 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 265 285 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 334 334 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 345 345 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 362 362 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 375 375 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 405 405 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 449 449 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 516 516 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 554 554 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 594 594 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 704 704 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 725 725 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 780 780 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 819 819 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 834 834 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 840 840 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 921 921 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1074 1074 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1200 1200 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1294 1294 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1311 1311 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1324 1324 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1336 1336 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1341 1341 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1358 1358 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1371 1371 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1449 AA; 159957 MW; 971BBAE191FDD1AF CRC64;

Query Match 43.5%; Score 60; DB 1; Length 1449;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
 Db 30 GHNWLIETFLNYSRLSP 49

RESULT 5
 ID VGL2_CVPFI STANDARD; PRT; 1449 AA.
 AC P33470;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

05-JUL-2004 (Rel. 44, Last annotation update)
E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
Name=S;
OS Porcine transmissible gastroenteritis coronavirus (strain Miller)
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=33737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91353366; PubMed=1966416;
RA Wesley R.D.;
RT "Nucleotide sequence of the E2-peplomer protein gene and partial
RT nucleotide sequence of the upstream polymerase gene of transmissible
RT gas gastroenteritis virus (Miller strain).";
RL Adv. Exp. Med. Biol. 276:301-306(1990).
CC !- FUNCTION: The peplomer protein mediates the binding of virions to
CC the host cell receptor and is involved in membrane fusion and in
CC syncytium formation.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S51223; AAB19567.2; -;
DR PIR; A43573; A43573.
DR InterPro; IPR002551; Corona_S1.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
KW Envelope protein; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 16
FT CHAIN 17 1449
FT DOMAIN 17 1390
FT TRANSMEM 1391 1410
FT DOMAIN 1411 1449
FT DOMAIN 1411 1432
FT CARBOHYD 26 26
FT CARBOHYD 42 42
FT CARBOHYD 71 71
FT CARBOHYD 94 94
FT CARBOHYD 243 243
FT CARBOHYD 250 250
FT CARBOHYD 285 285
FT CARBOHYD 334 334
FT CARBOHYD 345 345
FT CARBOHYD 362 362
FT CARBOHYD 375 375
FT CARBOHYD 405 405
FT CARBOHYD 449 449
FT CARBOHYD 516 516
FT CARBOHYD 532 532
FT CARBOHYD 554 554
FT CARBOHYD 562 562
FT CARBOHYD 584 584
FT CARBOHYD 704 704
FT CARBOHYD 725 725
FT CARBOHYD 780 780
FT CARBOHYD 819 819
FT CARBOHYD 834 834
FT CARBOHYD 840 840
FT CARBOHYD 921 921
FT CARBOHYD 1074 1074
FT CARBOHYD 1200 1200
FT CARBOHYD 1294 1294
FT CARBOHYD 1311 1311
FT CARBOHYD 1324 1324
FT CARBOHYD 1336 1336

FT CARBOHYD 1341 1341 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1358 1358 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1371 1371 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1449 AA; 159916 MW; PFC3CB8A4C9FB655 CRC64;
Query Match 43.5%; Score 60; DB 1; Length 1449;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNLYPLFLSP 24
DB 30 GHWNLITETFLNYSRLSP 49
RESULT 6
Q6PQ05 PRELIMINARY; PRT; 1449 AA.
AC Q6PQ05;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Spike glycoprotein.
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN2002;
RL He K., Yang Q.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY587882; AAT00645.1; -;
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1449 AA; 159953 MW; 237B61F2373933A1 CRC64;
Query Match 43.5%; Score 60; DB 2; Length 1449;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNLYPLFLSP 24
DB 30 GHWNLITETFLNYSRLSP 49
RESULT 7
Q7T428 PRELIMINARY; PRT; 1449 AA.
AC Q7T428;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Spike glycoprotein.
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TS;
RL Cheng J., Liu J.X., Wu R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY335548; AAQ02623.1; -;
DR EMBL; AY335549; AAQ02624.1; -;
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1449 AA; 159938 MW; 65D65801C851D0F5 CRC64;
Query Match 43.5%; Score 60; DB 2; Length 1449;

Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHWNLITFLLNYSRLSP 49

RESULT 8

AA000645 PRELIMINARY; PRT; 1449 AA.
AC AA000645; (Created)
DT 10-MAY-2004 (T-EMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (T-EMBLrel. 27, Last sequence update)
DE Spike glycoprotein.
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN2002;
RA He K., Yang Q.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY587882; AA000645.1; -
SQ SEQUENCE 1449 AA; 159953 MW; 237B61F2373933A1 CRC64;

Query Match 43.5%; Score 60; DB 2; Length 1449;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHWNLITFLLNYSRLSP 49

RESULT 9

Q9QR00 PRELIMINARY; PRT; 808 AA.
AC Q9QR00; (Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE S protein (Fragment).
OS Canine coronavirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11153;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UCD-1;
RX MEDLINE=99402069; PubMed=10475084;
RA Wesley R.D.;
RT "The S gene of canine coronavirus, strain UCD-1, is more closely related to the S gene of transmissible gastroenteritis virus than to that of feline infectious peritonitis virus.";
RL Virus Res. 61:145-152(1999).
DR EMBL; AF116248; AAD48907.1; -
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
FT NON_TER 808
SQ SEQUENCE 808 AA; 89556 MW; 5015CC8D4EA305C6 CRC64;

Query Match 42.8%; Score 59; DB 2; Length 808;
Best Local Similarity 55.0%; Pred. No. 9.8;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHWNLITFLLNYSRLSP 49

RESULT 10

O41204 PRELIMINARY; PRT; 184 AA.
AC O41204; (Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Glycoprotein S (Fragment).
GN Namespike;
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TMK-22;
RX MEDLINE=20128444; PubMed=1065059;
RA Aminev A.G., Amineva S.P., Baborenko E.P., Mishchenko V.A., Gusev A.A.;
RT "Restriction mapping of genome and nucleotide sequence of Aujeszky's disease virus strains isolated in CIS.";
RL Vopr. Virusol. 44:254-257(1999).
DR EMBL; Y15447; CAA75637.1; -
FT NON_TER 184
SQ SEQUENCE 184 AA; 21010 MW; 31E308C3B21D76D2 CRC64;

Query Match 39.9%; Score 55; DB 2; Length 184;
Best Local Similarity 55.0%; Pred. No. 8.4;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHWNLITFLLNYSRLSP 49

RESULT 11

Q8QRR7 PRELIMINARY; PRT; 789 AA.
AC Q8QRR7; (Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Spike glycoprotein (Fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HKT2;
RX MEDLINE=22699991; PubMed=12814887;
RA Kim S.J., Han J.H., Kwon H.M.;
RT "Partial sequence of the spike glycoprotein gene of transmissible gastroenteritis viruses isolated in Korea.";
RL Vet. Microbiol. 94:195-206(2003).
DR EMBL; AF481366; AAL89745.1; -
DR InterPro; IPR002551; Corona_S1.
DR Pfam; PF01600; Corona_S1; 1.
FT NON_TER 789
SQ SEQUENCE 789 AA; 87613 MW; 26BCF4FA9B461329 CRC64;

Query Match 39.9%; Score 55; DB 2; Length 789;
Best Local Similarity 55.0%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNHWNLITFLLNYSRLSP 49

RESULT 12

Q8QRR9
ID Q8QRR9 PRELIMINARY; PRT; 789 AA.
AC Q8QRR9
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Spike glycoprotein (fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT6;
RA Kim S.J., Kwon H.M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF481364; AAL89743.1; -;
DR InterPro; IPR002551; Corona S1.
DR Pfam; PF01600; Corona S1; 1.
DR NON_TER 789
FT NON_TER 789
SQ SEQUENCE 789 AA; 87647 MW; 17F00DB0EB92BB CRC64;
Query Match 39.9%; Score 55; DB 2; Length 789;
Best Local Similarity 55.0%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNEHNLIEFLNYSRLPP 49
RESULT 13
ID Q8QRR9 PRELIMINARY; PRT; 789 AA.
AC Q8QRR9
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Spike glycoprotein (fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT6;
RA Kim S.J., Kwon H.M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF481363; AAL89742.1; -;
DR InterPro; IPR002551; Corona S1.
DR Pfam; PF01600; Corona S1; 1.
DR NON_TER 789
FT NON_TER 789
SQ SEQUENCE 789 AA; 87496 MW; B941CF273FA296B0 CRC64;
Query Match 39.9%; Score 55; DB 2; Length 789;
Best Local Similarity 55.0%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNEHNLIEFLNYSRLPP 49
RESULT 14
ID Q8QRR9 PRELIMINARY; PRT; 789 AA.
AC Q8QRR9
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Spike glycoprotein (fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT2;
RA Kim S.J., Han J.H., Kwon H.M.;
RL MEDLINE=22699991; PubMed=12814887;
RT "Partial sequence of the spike glycoprotein gene of transmissible gastroenteritis viruses isolated in Korea."
RT Vet. Microbiol. 94:195-206(2003).
DR EMBL; AF481360; AAL89739.1; -;
DR InterPro; IPR002551; Corona S1.
DR Pfam; PF01600; Corona S1; 1.
DR NON_TER 789
FT NON_TER 789
SQ SEQUENCE 789 AA; 87600 MW; 615882B621F24FF2 CRC64;
Query Match 39.9%; Score 55; DB 2; Length 789;
Best Local Similarity 55.0%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNEHNLIEFLNYSRLPP 49
RESULT 15
ID Q8QRR9 PRELIMINARY; PRT; 790 AA.
AC Q8QRR9
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Spike glycoprotein (fragment).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus; Group 1 species.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT3;
RA Kim S.J., Han J.H., Kwon H.M.;
RL MEDLINE=22699991; PubMed=12814887;
RT "Partial sequence of the spike glycoprotein gene of transmissible gastroenteritis viruses isolated in Korea."
RT Vet. Microbiol. 94:195-206(2003).
DR EMBL; AF481361; AAL89740.1; -;
DR InterPro; IPR002551; Corona S1.
DR Pfam; PF01600; Corona S1; 1.
DR NON_TER 790
FT NON_TER 790
SQ SEQUENCE 790 AA; 87728 MW; 668351616011CED6 CRC64;
Query Match 39.9%; Score 55; DB 2; Length 790;
Best Local Similarity 55.0%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 5 GRHWGLPLSFLNYPFLSP 24
DB 30 GNEHNLIEFLNYSRLPP 49
Search completed: November 10, 2004, 14:50:12
Job time : 33.9698 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 14.6415 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-52

Perfect score: 194

Sequence: 1 ASWAPVGRDAETLQKQETIKAFKLKLEALMASNDNANKT 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	57	29.4	84	3	US-09-299-495F-11
2	57	29.4	88	2	US-08-690-011A-11
3	56	28.9	427	4	US-09-134-000C-4847
4	55	28.4	97	4	US-09-513-999C-5543
5	55	28.4	223	4	US-09-344-624-12
6	55	28.4	618	4	US-09-934-901-18
7	55	28.4	618	4	US-09-934-868-8
8	55	28.4	618	4	US-10-321-210-18
9	55	28.4	618	4	US-10-320-874-18
10	54	27.8	373	4	US-09-252-991A-28902
11	53	27.3	232	4	US-09-328-352-8025
12	53	27.3	263	4	US-09-328-352-7488
13	53	27.3	1086	4	US-09-543-681A-7696
14	52	26.8	235	4	US-09-620-405B-487
15	52	26.8	235	4	US-09-604-287A-487
16	52	26.8	235	4	US-09-834-753-487
17	52	26.8	285	4	US-09-886-319A-82
18	52	26.8	380	1	US-08-585-758A-4
19	52	26.8	380	1	US-08-977-818-4
20	52	26.8	380	2	US-08-670-274B-4
21	52	26.8	380	3	US-09-146-187-4
22	52	26.8	381	1	US-08-585-758A-2
23	52	26.8	381	1	US-08-977-818-2
24	52	26.8	381	2	US-08-670-274B-2
25	52	26.8	381	2	US-08-786-999-3
26	52	26.8	381	3	US-09-146-187-2
27	52	26.8	381	4	US-09-216-387-3

28	52	26.8	390	2	US-08-786-999-1	Sequence 1, Appli
29	52	26.8	390	4	US-09-216-387-1	Sequence 1, Appli
30	52	26.8	390	4	US-09-886-319A-2	Sequence 2, Appli
31	52	26.8	391	4	US-09-886-319A-1	Sequence 1, Appli
32	52	26.8	505	4	US-09-620-405B-478	Sequence 478, App
33	52	26.8	505	4	US-09-620-405B-485	Sequence 485, App
34	52	26.8	505	4	US-09-604-287A-478	Sequence 478, App
35	52	26.8	505	4	US-09-604-287A-485	Sequence 485, App
36	52	26.8	505	4	US-09-834-759-478	Sequence 478, App
37	52	26.8	505	4	US-09-834-759-485	Sequence 485, App
38	52	26.8	505	4	US-09-590-751A-478	Sequence 478, App
39	52	26.8	608	4	US-09-328-352-5057	Sequence 5057, Ap
40	51.5	26.5	534	4	US-09-103-664A-2	Sequence 2, Appli
41	51	26.3	90	4	US-09-513-999C-6463	Sequence 6463, Ap
42	51	26.3	588	1	US-07-903-466-3	Sequence 3, Appli
43	51	26.3	588	5	PCT-US93-05294-3	Sequence 3, Appli
44	50	25.8	151	4	US-09-270-767-47583	Sequence 47583, A
45	50	25.8	1151	4	US-09-710-279-2448	Sequence 2448, Ap

ALIGNMENTS

RESULT 1

US-09-299-495F-11

; Sequence 11, Application US/09299495F

; Patent No. 6361968

; GENERAL INFORMATION:

; APPLICANT: VINSON, Charles R.

; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

; CELLULAR PROTEIN

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0053

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS WORD 97

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/299,495F

; FILING DATE: 26-APR-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/001,654

; FILING DATE: 31-JUL-1995

; APPLICATION NUMBER: 60/018,496

; FILING DATE: 29-MAY-1996

; APPLICATION NUMBER: 08/690,011

; FILING DATE: 31-JULY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Serunian, Leslie A.

; REGISTRATION NUMBER: 35,353

; REFERENCE/DOCKET NUMBER: 2026-4199US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)758-4800

; TELEFAX: (212)751-6849

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 84 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-299-495F-11

```

Query Match      29.4%; Score 57; DB 3; Length 84;
Best Local Similarity 37.1%; Pred. No. 2.7;
Matches 13; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy   6 VGRDAETLQKOKETIKAFKLKKLEALMASNDNANKT 40
     :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db   33 LARENEELEKEAELEQLAELNRVAVLENQNKT 67

RESULT 2
US-08-690-011A-11
; Sequence 11, Application US/086900011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199U51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO. 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-690-011A-11

Query Match      29.4%; Score 57; DB 2; Length 88;
Best Local Similarity 37.1%; Pred. No. 2.8;
Matches 13; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy   6 VGRDAETLQKOKETIKAFKLKKLEALMASNDNANKT 40
     :|:~::~::||:|:|:|:|:|:|:|:|:|
Db   33 LARENEELEKEAELEQLAELNRVAVLENQNKT 67

RESULT 3
US-09-134-000C-4847
; Sequence 4847, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

```



```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-934-624-12

Query Match      28.4%; Score 55; DB 4; Length 223;
Best Local Similarity 42.1%; Pred. No. 14;
Matches 16; Conservative 5; Mismatches 11; Indels 6; Gaps 1;

QY 8 RDAETLQKQKETIKA-----FLKKLEALMASNDNANK 39
DB 65 KDSLNGKRPQYIKAKENTSHKIKKLEAAKNSLQNAOK 102

RESULT 6
US-09-934-901-18
; Sequence 18, Application US/09934901
; Patent No. 655353
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 655353ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-18

Query Match      28.4%; Score 55; DB 4; Length 618;
Best Local Similarity 34.5%; Pred. No. 44;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 11 ETLOKQKETIKAFLLKKLEALMASNDNANK 39
DB 23 QVIARSRETRAAYLKRIEAAIAEGPQRNK 51

RESULT 7
US-09-934-868-8
; Sequence 8, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-868-8

Query Match      28.4%; Score 55; DB 4; Length 618;
Best Local Similarity 34.5%; Pred. No. 44;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 11 ETLOKQKETIKAFLLKKLEALMASNDNANK 39

```

```

DB 23 QVIARSRETRAAYLKRIEAAIAEGPQRNK 51

RESULT 8
US-10-321-210-18
; Sequence 18, Application US/10321210
; Patent No. 6767744
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6767744ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-18

Query Match      28.4%; Score 55; DB 4; Length 618;
Best Local Similarity 34.5%; Pred. No. 44;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 11 ETLOKQKETIKAFLLKKLEALMASNDNANK 39
DB 23 QVIARSRETRAAYLKRIEAAIAEGPQRNK 51

RESULT 9
US-10-320-874-18
; Sequence 18, Application US/10320874
; Patent No. 6773905
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6773905ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,874
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-320-874-18

Query Match      28.4%; Score 55; DB 4; Length 618;
Best Local Similarity 34.5%; Pred. No. 44;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 11 ETLOKQKETIKAFLLKKLEALMASNDNANK 39
DB 23 QVIARSRETRAAYLKRIEAAIAEGPQRNK 51

RESULT 10
US-09-252-991A-28902

```

```
; Sequence 28902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28902
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28902

Query Match      27.8%; Score 54; DB 4; Length 373;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 12; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY      1 ASMAPVGRDAETLQKQETIKAFKLEALMAS 33
      : : : : : : : : : : : : : : : : : :
Db      193 SOLAKARADLEAQRVQDRTLGAIEIKRLBAQIAS 225

RESULT 11
US-09-328-352-8025
; Sequence 8025, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8025
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8025

Query Match      27.3%; Score 53; DB 4; Length 232;
Best Local Similarity 42.9%; Pred. No. 27;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY      1 ASMAPVGRDAETLQKQETIKAFKLE 28
      : : : : : : : : : : : : : : : :
Db      174 AIGQVGFDAAIRQNVETLVADLKKL 201

RESULT 12
US-09-328-352-7488
; Sequence 7488, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7488
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7488
```

```
Query Match      27.3%; Score 53; DB 4; Length 263;
Best Local Similarity 39.5%; Pred. No. 31;
Matches 17; Conservative 3; Mismatches 13; Indels 10; Gaps 2;

QY      6 VGRDA-----ETLQKQETIKAFKLEALMASND----NAN 38
      : : : : : : : : : : : : : : : : : :
Db      66 VGIDADAEQLVEVQKQKTDALDAVNSLNRAVQERDVAVTNAN 108

RESULT 13
US-09-543-681A-7696
; Sequence 7696, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7696
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7696

Query Match      27.3%; Score 53; DB 4; Length 1086;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 16; Conservative 5; Mismatches 15; Indels 2; Gaps 1;

QY      4 APVGRDAETLQKQETIKAFKLEALMASND--NANK 39
      : : : : : : : : : : : : : : : : : :
Db      472 ASIGKLAEGLDIAQIEKIBELKRRQDLKSSDLGGANK 509

RESULT 14
US-09-620-405B-487
; Sequence 487, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 487
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-620-405B-487

Query Match      26.8%; Score 52; DB 4; Length 235;
Best Local Similarity 38.2%; Pred. No. 37;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY      4 APVGRDAETLQKQETIKAFKLEALMASNDNA 37
      : : : : : : : : : : : : : : : :
Db      36 ATVIRHGETLRETKEEINELNENVIQRLTAEVENA 69

RESULT 15
```

Fri Nov 12 14:55:37 2004

```

US-09-604-287A-487
; Sequence 487, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yudiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 487
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-287A-487

```

```

Query Match      26.8%; Score 52; DB 4; Length 235;
Best Local Similarity 38.2%; Pred. No. 37;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

```

```

QY      4. APVGRDAETLQKQETIKAFLLKLEALMASNDNA 37
DB      36 ATVIRHGETLRRTKKEINELNRMIOQLTAEVENA 69

```

```

Search completed: November 10, 2004, 14:55:38
Job time : 15.6915 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 46.2642 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-52

Perfect score: 194

Sequence: 1 ASMAPVGRDAETLQKQETIKAFKLEALMASNDNANKT 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35325886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	194	100.0	40	14	US-10-092-750-52
2	190	97.9	2835	9	US-09-885-535-4
3	190	97.9	5171	16	US-10-408-765A-2687
4	186	95.9	1174	14	US-10-094-749-1911
5	67	34.5	490	14	US-10-029-386-33994
6	67	34.5	5373	16	US-10-408-765A-741
7	66	34.0	990	15	US-10-108-260A-3896
8	62.5	32.2	174	15	US-10-282-122A-47358
9	62.5	32.2	340	15	US-10-282-122A-52947
10	60	30.9	103	17	US-10-425-115-236832
11	57	29.4	84	14	US-10-059-720-11
12	55	28.4	217	17	US-10-425-115-351517
13	55	28.4	240	10	US-09-764-891-2917

```

55 28.4 618 9 US-09-934-901-18 Sequence 18, Appl
55 28.4 618 9 US-09-934-868-8 Sequence 8, Appl
55 28.4 618 14 US-10-320-924-18 Sequence 18, Appl
55 28.4 618 14 US-10-320-874-18 Sequence 18, Appl
55 28.4 618 16 US-10-321-210-18 Sequence 18, Appl
55 28.4 1225 16 US-10-408-765A-166 Sequence 166, App
55 28.4 3672 14 US-10-369-493-6146 Sequence 6146, App
55 28.4 4101 14 US-10-369-493-6509 Sequence 6509, App
55 28.4 4101 14 US-10-369-493-6510 Sequence 6510, App
54 27.8 135 15 US-10-424-599-199232 Sequence 199232, A
54 27.8 171 15 US-10-425-114-36706 Sequence 36706, A
54 27.8 180 16 US-10-437-963-151123 Sequence 151123, A
54 27.8 213 17 US-10-767-701-54358 Sequence 54358, A
54 27.8 741 15 US-10-425-115-336495 Sequence 336495, A
54 27.8 772 17 US-10-425-115-267069 Sequence 267069, A
54 27.8 1197 17 US-10-425-115-190351 Sequence 190351, A
54 27.8 3562 14 US-10-341-434-109 Sequence 109, App
53 27.6 112 15 US-10-264-237-1771 Sequence 1771, App
53 27.6 583 15 US-10-282-122A-61644 Sequence 61644, A
53 27.6 3551 15 US-10-263-939-144 Sequence 144, App
53 27.6 3664 14 US-10-177-293-423 Sequence 423, App
53 27.6 3664 15 US-10-263-929-2287 Sequence 2287, App
53 27.6 3664 16 US-10-408-765A-2287 Sequence 205253, A
53 27.3 70 17 US-10-425-115-217430 Sequence 217430, A
53 27.3 107 17 US-10-425-115-217430 Sequence 1, Appl
53 27.3 228 13 US-10-013-379-1 Sequence 45109, A
53 27.3 231 15 US-10-282-122A-45109 Sequence 36097, A
53 27.3 288 16 US-10-767-701-36097 Sequence 23, Appl
53 27.3 685 17 US-10-416-330-23 Sequence 175207, A
53 27.3 687 16 US-10-437-963-175207 Sequence 64, Appl
53 27.3 933 10 US-09-882-227-64

```

ALIGNMENTS

```

RESULT 1
US-10-092-750-52
; Sequence 52, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION: Philip W.
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-52

```

```

Query Match 100.0%; Score 194; DB 14; Length 40;
Best Local Similarity 100.0%; Pred. No. 5e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ASMAPVGRDAETLQKQETIKAFKLEALMASNDNANKT 40
Db 1 ASMAPVGRDAETLQKQETIKAFKLEALMASNDNANKT 40

```

```

RESULT 2
US-09-885-535-4
; Sequence 4, Application US/09885535
; Patent No. US20020104105A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Reichman, Karen
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-266-II
; CURRENT APPLICATION NUMBER: US/09/885,535
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,245
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-885-535-4

Query Match          97.9%; Score 190; DB 9; Length 2835;
Best Local Similarity 100.0%; Pred. No. 2.1e-15;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  SNAPVGRDAETLQKQKETIKAFKLEALMASNDNANKT 40
Db      557  SNAPVGRDAETLQKQKETIKAFKLEALMASNDNANKT 595

RESULT 3
US-10-408-765A-2687
; Sequence 2687, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2687
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 4471
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-2687

Query Match          97.9%; Score 190; DB 16; Length 5171;
Best Local Similarity 100.0%; Pred. No. 4.3e-15;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  SNAPVGRDAETLQKQKETIKAFKLEALMASNDNANKT 40
Db      2784  SNAPVGRDAETLQKQKETIKAFKLEALMASNDNANKT 2822

RESULT 4
US-10-094-749-1911
; Sequence 1911, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI

```

```

; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1911
; LENGTH: 1174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1911

Query Match          95.3%; Score 186; DB 14; Length 1174;
Best Local Similarity 97.4%; Pred. No. 2.6e-15;
Matches 38; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2  SNAPVGRDAETLQKQKETIKAFKLEALMASNDNANKT 40
Db      343  SNAPVGRDAETLQKQKETIKAFKLEALIASNDNANKT 381

RESULT 5
US-10-029-386-33994
; Sequence 33994, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33994
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL137853.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.3
; OTHER INFORMATION: SWISSPROT HIT: Q9UPN3, EVALUATION 0.00e+00
US-10-029-386-33994

Query Match          34.5%; Score 67; DB 14; Length 490;
Best Local Similarity 42.3%; Pred. No. 3.3;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY      3  MAPVGRDAETLQKQKETIKAFKLEAL 30
Db      363  MGAIGRDTSLQSQIEDVRLFNKHL 390

```

```

RESULT 6
US-10-408-765A-741
; Sequence 741, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 741
; LENGTH: 5373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-741

Query Match      34.5%; Score 67; DB 16; Length 5373;
Best Local Similarity 42.9%; Pred. No. 51;
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY   3 MAPVGRDAETLQKQKETIKAFLLKKLEAL 30
    |::|::|||::|::|::|::|::|::|::|
DB   3027 MGAIGRDTLSQSIEDVRFLFNKHIVL 3054

RESULT 7
US-10-108-260A-3896
; Sequence 3896, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3896
; LENGTH: 990
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3896

Query Match      34.0%; Score 66; DB 15; Length 990;
Best Local Similarity 51.7%; Pred. No. 9.9;
Matches 15; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY   1 ASMAPVGRDAETLQKQKETIKAFLLKKLEA 29
    |||::|::|::|::|::|::|::|::|
DB   713 ASMSPIGTDLTVRKQQIEELKQF--KSEA 739

RESULT 8
US-10-282-122A-47358
; Sequence 47358, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; REMAINING PRIOR APPLICATION DATA REMOVED - SEE FILE WRAPPER OR PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47358
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47358

Query Match      32.2%; Score 62.5; DB 15; Length 174;
Best Local Similarity 42.1%; Pred. No. 3.9;
Matches 16; Conservative 6; Mismatches 13; Indels 3; Gaps

QY   5 PVGRDAETLQKQKETIK---APFKKLEALMASNDNVANK 39
    |||::|::|::|::|::|::|::|::|
DB   133 PDERSREELLKENEYLAEAVAYLKLDALLAKQAQK 170

RESULT 9
US-10-282-122A-52947
; Sequence 52947, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 52947
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Clostridium difficile
; US-10-282-122A-52947

Query Match      32.2%; Score 62.5; DB 15; Length 340;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 18; Conservative 7; Mismatches 8; Indels 3; Gaps 2;

Qy      6 VGRDAETLQKQETIKAF--LKKLALMA-SNDNAN 38
Db      288 VGIDETVQVKETIAFAGTEKVEALLAISNHQKN 323

RESULT 10
US-10-425-115-236832
; Sequence 236832, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 236832
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147579C.1.pep
; US-10-425-115-236832

Query Match      30.9%; Score 60; DB 17; Length 103;
Best Local Similarity 41.4%; Pred. No. 4.5;
Matches 12; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy      2 SMAPVGRDAETLQKQETIKAFLEKLEAL 30
Db      30 STAPVDRDSEVTPEDRQVQIYVHDEAV 58

RESULT 11
US-10-059-720-11
; Sequence 11, Application US/10059720
; Publication No. US20030027314A1

; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
; US-10-059-720-11

Query Match      29.4%; Score 57; DB 14; Length 84;
Best Local Similarity 37.1%; Pred. No. 8.9;
Matches 13; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy      6 VGRDAETLQKQETIKAFLEKLEALMASNDNANKT 40
Db      33 LARENELEKEAELEQELAELENRVAVLEQNKT 57

RESULT 12
US-10-425-115-351517
; Sequence 351517, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 351517
; LENGTH: 217
; TYPE: PRT
```



```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFR4577_8374C.1.pep
US-10-425-115-351517

Query Match      28.4%; Score 55; DB 17; Length 217;
Best Local Similarity 46.2%; Pred. No. 48;
Matches 12; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 9 DAETLQKQKETIKAFKKLEALMASN 34
Db 89 DLKLNKNKLIKLLAKKYDAFMASD 114

RESULT 13
US-09-764-891-2917
; Sequence 2917, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2917
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (220)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (236)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-2917

Query Match      28.4%; Score 55; DB 10; Length 240;
Best Local Similarity 36.4%; Pred. No. 54;
Matches 12; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 2 SNAPVGRDAETLQKQKETIKAFKKLEALMASN 34
Db 128 SMKAIQKQCTLQKNKKLKQEVVNLKSYMERN 160

RESULT 14
US-09-934-901-18
; Sequence 18, Application US/09934901
; Patent No. US20020110885A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. US20020110885A1ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934,901
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-18

Query Match      28.4%; Score 55; DB 9; Length 618;
```

```
Best Local Similarity 34.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 11 ETLOKQKETIKAFKKLEALMASNDNANK 39
Db 23 QVIARSRETRAAYLKIEIAAIEAGPQRNK 51

RESULT 15
US-09-934-868-8
; Sequence 8, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 618
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-868-8

Query Match      28.4%; Score 55; DB 9; Length 618;
Best Local Similarity 34.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 11 ETLOKQKETIKAFKKLEALMASNDNANK 39
Db 23 QVIARSRETRAAYLKIEIAAIEAGPQRNK 51

Search completed: November 11, 2004, 02:43:08
Job time : 47.3142 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 9.81132 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-52

Perfect score: 194

Sequence: 1 ASMAPVGRDAETLQKQETIKAFLLKLEALMASNDANKT 40

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	30.4	83	2 I48283	gene CC2 protein -
2	57.5	29.6	554	2 C82743	ABC transporter AT
3	57	29.4	247	2 F84004	hypothetical prote
4	57	29.4	318	2 G82416	probable resolvas
5	57	28.4	419	2 A35438	keratin, type I cy
6	56.5	29.1	447	2 F75164	primase dnag-like
7	56.5	29.1	447	2 E71177	hypothetical prote
8	56	28.9	299	2 C80519	hypothetical prote
9	56	28.9	3660	1 S02041	dystrophin, muscle
10	55	28.4	480	2 A56694	keratin, type II,
11	55	28.4	1065	2 T25068	hypothetical prote
12	55	28.4	1233	2 I43483	chromosome segrega
13	55	28.4	1233	2 T30534	chromosome segrega
14	55	28.4	3672	2 T23433	hypothetical prote
15	55	28.4	3704	2 T37316	probable laminin a
16	55	28.4	4063	2 T42993	probable spectrin
17	55	28.4	4101	2 T23630	hypothetical prote
18	54.5	28.1	4151	2 T13734	groovin gene prote
19	54	27.8	173	2 T21815	hypothetical prote
20	54	27.8	355	2 G83254	probable secretion
21	54	27.8	473	2 D81322	outer membrane pro
22	54	27.8	3685	1 A27605	dystrophin, muscle
23	53	27.3	190	2 S70261	outer surface prot
24	53	27.3	209	2 I40281	outer surface prot
25	53	27.3	229	2 S66577	ribosomal protein
26	53	27.3	527	2 F85378	conserved hypotnet
27	53	27.3	933	2 E84603	hypothetical prote
28	53	27.3	1042	2 G84514	type I restriction
29	53	27.3	1188	2 T19552	hypothetical prote

30	53	27.3	3678	2 S28916	dystrophin - mouse
31	52.5	27.1	173	2 S55150	hypothetical prote
32	52	26.8	257	2 I38025	keratin-like prote
33	52	26.8	274	2 T13010	hypothetical prote
34	52	26.8	377	2 D72317	hypothetical prote
35	52	26.8	398	2 T04501	hypothetical prote
36	52	26.8	915	2 T26695	hypothetical prote
37	52	26.8	1244	2 T23744	chromosome-associa
38	52	26.8	1274	2 T02636	D1 protein homolog
39	52	26.8	3450	2 T26963	hypothetical prote
40	52	26.8	3461	2 T26964	hypothetical prote
41	51.5	26.5	232	2 T23040	hypothetical prote
42	51.5	26.5	279	2 E97138	hypothetical prote
43	51.5	26.5	425	2 E64403	translation releas
44	51	26.3	233	2 A82132	conserved hypotnet
45	51	26.3	286	2 T19460	hypothetical prote

ALIGNMENTS

RESULT 1

I48283

gene CC2 protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: I48283

R:Mauchner, A.; Camonis, J.H.; Sobel, A.

Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995

A>Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dc

A:Reference number: I48282; MUID:95241452; PMID:7724523

A:Accession: I48283

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-83 <RES>

A:Cross-references: EMBL:X82319; NID:g791075; PIDN:CAA57762.1; PID:g791076

C:Genetics:

A:Gene: CC2

Query Match 30.4%; Score 59; DS 2; Length 83;
Best Local Similarity 36.4%; Pred. No. 2.4;
Matches 12; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 3 MAPVGRDAETLQKQETIKAFLLKLEALMASND 35

Db 41 VAEVGNKIELLKKKDELSALEKMEHQSENND 73

RESULT 2

C82743

ABC transporter ATP-binding protein XF0944 [imported] - Xylella fastidiosa (strain 9a5c

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2004

C:Accession: C82743

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ

Nature 406, 151-157, 2000

A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: C82743

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-554 <Sim>

A:Cross-references: UNIPROT:Q9PFT4; GB:AE003933; GB:AE003849; NID:g9105863; PIDN:AAF83;

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canarog, L.E.A.; Carraro, D.M.; Carter,

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Lal

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.N.; Miracca, E.C.; Miyaki, C.Y.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshahko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
A;Content: annotation
C;Genetics:
A;Gene: XP0944
C;Superfamily: ATP-binding cassette homology

Query Match 29.6%; Score 57.5; DB 2; Length 554;
Best Local Similarity 40.0%; Pred. No. 28;
Matches 14; Conservative 7; Mismatches 7; Indels 7; Gaps 1;

Qy 1 ASWAPGRDAETLQOKETIKAFKLEALMASND 35
Db 110 AAABEGADFDALAKEQE-----RLEAILASND 137

RESULT 3
F84004
hypothetical protein BH2838 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: F84004
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F84004
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <STO>
A;Cross-references: UNIPROT:Q9K913; GB:AP001516; GB:BA000004; NID:gi0175192; PIDN:BA065 A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2838

Query Match 29.4%; Score 57; DB 2; Length 247;
Best Local Similarity 38.7%; Pred. No. 14;
Matches 12; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy 8 RDAETLQOKETIKAFKLEALMASNDNAN 38
Db 36 QQAELVQKQTIALEKLVQQLQEVNEAAS 66

RESULT 4
G82416
probable resolvase VCA0795 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82416
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82416
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <HEI>
A;Cross-references: UNIPROT:Q9KLF1; GB:AE004407; GB:AE003853; NID:g9658210; PIDN:AAF9669 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0795
A;Map position: 2

Query Match 29.4%; Score 57; DB 2; Length 318;
Best Local Similarity 39.4%; Pred. No. 18;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 4 APVGRDAETLQOKETIKAFKLEALMASNDN 36
Db 75 ARIGGHCELLKGAKKIKALEKKADSFIPSDS 107

RESULT 5
A25438
keratin, type I cytoskeletal protein, B2 - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: A25438
R;Miyatani, S.; Winkles, J.A.; Sargent, T.D.; Dawid, I.B.
J. Cell Biol. 103, 1957-1965, 1986
A;Title: Stage-specific keratins in Xenopus laevis embryos and tadpoles: the XK81 gene A;Reference number: A25438; MUID:87057649; PMID:2430981
A;Accession: A25438
A;Molecule type: mRNA
A;Residues: 1-419 <MIY>
A;Cross-references: UNIPROT:P05781; GB:M18155; GB:X05865; NID:g214558; PID:g214560 C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 29.4%; Score 57; DB 2; Length 419;
Best Local Similarity 35.7%; Pred. No. 24;
Matches 10; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy 7 GRDAETLQOKETIKAFKLEALMASN 34
Db 80 GNDKQTMQNLDRSLASYLEKVRALFAAN 107

RESULT 6
F75164
primase dnag-like PAB0316 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75164
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str A;Reference number: A75001
A;Accession: F75164
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <KAW>
A;Cross-references: UNIPROT:Q9V1F2; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB493 A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB0316
C;Superfamily: conserved hypothetical protein MJ1206

Query Match 29.1%; Score 56.5; DB 2; Length 447;
Best Local Similarity 44.4%; Pred. No. 29;
Matches 16; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

Qy 4 APVGRDAETLQOKETIKAFKLEALMASNDNAN 39
Db 272 APBGKEVEELTK-KEIKALRSKVPAPAEVYNELFNK 306

RESULT 7
E71177
hypothetical protein PH1699 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C;Accession: E71177
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71177
A;Status: preliminary; nucleic acid sequence not shown; translation not shown


```

Best Local Similarity 38.2%; Pred. No. 48;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 4 APVGRDAETLOKQKETIKAFKLKLEALMASNDNA 37
DB 323 ATVNRHGETLRTKKEINELNLIQLTAIENA 356

RESULT 11
T25068
hypothetical protein T21C9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25068
R;McMurray, A.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19977
A;Accession: T25068
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-1065 <WIL>
A;Cross-references: UNIPROT:Q22639; EMBL:Z73098; PIDN:CAA97331.1; GSPDB:GN000023; CESP:T2
A;Experimental source: clone T21C9
C;Genetics:
A;Gene: CESP:T21C9.2
A;Map position: 5
A;Introns: 67/2; 147/2; 207/3; 242/3; 293/3; 825/2; 897/1; 1032/2

Query Match 28.4%; Score 55; DB 2; Length 1065;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 8 RDAETLOKQKETIKAFKLKLEALMA 32
DB 133 RDTPIEEKKFTUHSFTQNSAVLA 157

RESULT 12
I54383
chromosome segregation protein smc1 [similarity] - human
N;Alternate names: protein SBI.8/DXS423E; structural maintenance of chromosomes (SMC) pr
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 02-Jun-2000
C;Accession: I54383
R;Rocques, P.J.; Clark, J.; Ball, S.; Crew, J.; Gill, S.; Christodoulou, Z.; Borts, R.H.
Hum. Mol. Genet. 4, 243-249, 1995
A;Title: The human SBI.8 gene (DXS423E) encodes a putative chromosome segregation protei
A;Reference number: I54383; MUID:195276737; PMID:7757074
A;Accession: I54383
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-1233 <RES>
A;Cross-references: GB:S78271; NID:G999379; PIDN:AAB34405.1; PID:G999380
C;Genetics:
A;Gene: GDB:X1A0178; SMC1; SBI.8/DXS423E
A;Cross-references: GDB:9785049
A;Map position: Xpter-Xqter
C;Superfamily: chromosome segregation protein SMC1

Query Match 28.4%; Score 55; DB 2; Length 1233;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 5; Mismatches 11; Indels 5; Gaps 1;

QY 8 RDAETLOKQKETIKA-----FLKKLEALMASNDNANK 39
DB 287 KDSLNQKRPQYIKAKENTSHKIKKLEAAKSLQNAQK 324

RESULT 13
T30534
chromosome segregation protein SMC1 homolog - Japanese pufferfish
C;Species: Fugu rubripes (Japanese pufferfish)
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

```

```

C;Accession: T30534
R;Riboldi Tunicliffe, G.R.; Platzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Rosei
submitted to the EMBL Data Library, September 1997
A;Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disea
A;Reference number: Z20848
A;Accession: T30534
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1233 <RIB>
A;Cross-references: UNIPROT:O73696; EMBL:AF026198; NID:G3098263; PID:G3098266; PIDN:AAC
C;Superfamily: chromosome segregation protein SMC1

Query Match 28.4%; Score 55; DB 2; Length 1233;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 5; Mismatches 11; Indels 6; Gaps 1;

QY 8 RDAETLOKQKETIKA-----FLKKLEALMASNDNANK 39
DB 287 KDSLNQKRPQYIKAKENTSHKIKKLEAAKSLQNAQK 324

RESULT 14
T23433
hypothetical protein K08C7.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23433
R;Berks, M.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19740
A;Accession: T23433
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-3672 <WIL>
A;Cross-references: UNIPROT:Q21313; EMBL:Z70286; PIDN:CAA94293.1; GSPDB:GN000022; CESP:K
C;Experimental source: clone K08C7
C;Genetics:
A;Gene: CESP:K08C7.3
A;Map position: 4
A;Introns: 66/1; 284/3; 563/1; 1187/3; 1248/3; 1300/1; 1460/1; 1623/3; 2361/3; 2988/3;
C;Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match 28.4%; Score 55; DB 2; Length 3672;
Best Local Similarity 38.2%; Pred. No. 4.2e+02;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 6 VGRDAETLOKQKETIKAFKLKLEALMASNDNANK 39
DB 2314 IQBETKLDKQKTFEAKQKRAELAAAYLNSAQQ 2347

RESULT 15
T37316
probable laminin alpha chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37316
R;Joh, K.; Zhu, K.; Hedgecock, E.M.; Inoue, T.; Hori, K.
submitted to the EMBL Data Library, August 1998
A;Description: Laminin alpha chain gene in the nematode C. elegans.
A;Reference number: Z21681
A;Accession: T37316
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-3704 <JOH>
A;Cross-references: UNIPROT:P91904; EMBL:AB016806; PIDN:BAA32347.1
A;Experimental source: strain N2
C;Genetics:
A;Gene: epi-1
A;Map position: IV
A;Introns: 66/1; 284/3; 563/1; 1187/3; 1248/3; 1300/1; 1460/1; 1623/3; 2361/3; 2988/3;
C;Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

```

Query Match 28.4%; Score 55; DB 2; Length 3704;
Best Local Similarity 38.2%; Pred. No. 4.2e+02;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 6 VGRDAETLOKOKETIKAFLLKKLEALMASNDNANK 39
Db 2314 IQEETEKLDKQXETFEAQKKAELAAVLSAQQ 2347

Search completed: November 10, 2004, 14:52:21
Job time : 10.8113 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 53.283 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-52

Perfect score: 194

Sequence: 1 ASMAPVGRDAETLQKQETIKAFLEKLEALMASNDNANKT 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	97.9	5171	1	BPEA_HUMAN
2	151	77.8	1589	2	Q6ZQ59
3	151	77.8	1589	2	BAC98011
4	151	77.8	1589	2	BPA1_MOUSE
5	73	37.6	5327	1	MACF_MOUSE
6	69	35.6	486	2	Q89QJ6
7	67	34.5	4433	2	BAA32310
8	67	34.5	5430	1	MACF_HUMAN
9	67	34.5	5938	1	MAC4_HUMAN
10	66	34.0	482	2	Q6N8H1
11	66	34.0	482	2	CAB27373
12	66	34.0	649	2	Q9LVQ4
13	65.5	33.8	286	2	Q6JIP4
14	65.5	33.8	1679	1	GCC2_MOUSE
15	65	33.5	636	2	Q82X23
16	61.5	31.7	469	2	Q73AL3
17	61.5	31.7	469	2	AAS40688
18	60.5	31.2	570	2	Q8STN2
19	60	30.9	776	2	Q8BN37
20	59.5	30.7	456	2	Q6HKR1
21	59	30.4	83	2	Q6LBE4
22	59	30.4	83	2	CAA57762
23	58	29.9	574	2	Q7P7A7
24	58	29.9	579	2	Q8RI43
25	58	29.9	1232	2	Q7ZTJ9
26	58	29.9	1669	1	DMDB_DROME
27	58	29.9	3497	1	DMDA_DROME
28	57.5	29.6	510	2	Q7YKQ3
29	57.5	29.6	514	2	Q8SM17
30	57.5	29.6	554	2	Q87AS1
31	57.5	29.6	554	2	Q9PET4

32 57 29.4 247 2 Q9K913
33 57 29.4 280 1 K1C3_XENLA
34 57 29.4 297 2 Q6N2T7
35 57 29.4 297 2 CAB29402
36 57 29.4 318 1 Q9KLF1
37 57 29.4 419 1 K1C4_XENLA
38 57 29.4 469 2 O7ZY83
39 57 29.4 693 1 EX70_DROME
40 57 29.4 900 2 Q7R6H3
41 56.5 29.1 317 2 Q9VT70
42 56.5 29.1 447 1 Y475_PYRAB
43 56.5 29.1 447 1 YG99_PYRHO
44 56.5 29.1 509 1 MATK_VERTH
45 56.5 29.1 511 2 Q7YKQ0

Q9K913 bacillus ha
P05782 xenopus lae
Q6N2T7 rhodopseu
CAB29402 rhodopseu
Q9KLF1 vibrio chol
P05781 xenopus lae
Q7ZY83 xenopus lae
Q9VJ8 drosophila
Q7R6H3 giardia lam
Q9VT70 drosophila
Y475 pyrocooc
YG99 pyrocooc
MATK verbasum t
Q7YKQ0 lamium macu

ALIGNMENTS

RESULT 1
BPEA_HUMAN
ID BPEA_HUMAN STANDARD; PRT; 5171 AA.
AC O94833; Q8N1T8; Q8N8J3; Q8WXX9; Q96AK9; Q96D05; Q9H555;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Bullous pemphigoid antigen 1, isoforms 6/9/10 (Trabeculin-beta)
DE (Bullous pemphigoid antigen) (BPA) (Hemidesmosomal plaque protein)
DE (Dystonia musculorum protein) (Dystonin).
GN Name-DST; Synonyms-BPAG1, DMH, DT, KIAA0728;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 6), AND TISSUE SPECIFICITY.
RC TISSUE=Keratinocytes;
RX MEDLINE=21839111; PubMed=11751855; DOI=10.1074/jbc.M109209200;
RA Okumura M., Yamakawa H., Ohara O., Owaribe K.;
RT "Novel alternative splicings of BPAG1 (bullous pemphigoid antigen 1) including the domain structure closely related to VACF (microtubule actin cross-linking factor).";
RT J. Biol. Chem. 277:6682-6687(2002).
RL [2]
RN SEQUENCE OF 1342-5171 FROM N.A. (ISOFORM 10).
RP TISSUE=Ductenium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Pzange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 1702-4156 FROM N.A. (ISOFORM 6).
RC TISSUE=Brain, Placenta, and Tongue;
RX PubMed=14702039; DOI=10.1038/ngi285;
RA Ora T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,

RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawal Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosori T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Sanoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4].
RP SEQUENCE OF 3535-5171 FROM N.A. (ISOFORM 6).
RC TISSUE=Brain;
RX MEDLINE=22158633; PubMed=1216954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.,
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
RN [5].
RP SEQUENCE OF 4107-5171 FROM N.A. (ISOFORM 6).
RC TISSUE=Brain;
RX MEDLINE=99087487; PubMed=9872452;
RA Tanaka A., Kotani H., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
RN [6].
RP SEQUENCE OF 4031-5171 FROM N.A. (ISOFORM 9).
RA Smith M.;
RN Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [7].
RP DISEASE.
RC TISSUE=Keratinocytes;
RX MEDLINE=92011493; PubMed=1717441;
RA Sawamura D., Li K., Chu M.-L., Uitto J.,
RT "Human bullous pemphigoid antigen (BPAG1). Amino acid sequences
deduced from cloned cDNAs predict biologically important peptide
segments and protein domains.";
RL J. Biol. Chem. 266:17784-17790(1991).
RN [7].
RP -!- FUNCTION: Cytoskeletal linker protein. Anchors keratin-containing
intermediate filaments to the inner plaque of hemidesmosomes. The
proteins may self-aggregate to form filaments or a two-dimensional
mesh (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Name=6; Synonyms=EA;
CC IsoId=O94833-2; Sequence=Displayed;
CC Name=1;
CC IsoId=Q03001-1; Sequence=External;
CC Name=2;
CC IsoId=Q03001-2; Sequence=External;
CC Name=3; Synonyms=1e; Sequence=External;
CC IsoId=Q03001-3; Sequence=External;
CC Name=4;
CC IsoId=Q03001-4; Sequence=External;
CC Name=5;
CC IsoId=Q03001-5; Sequence=External;
CC Name=6; Synonyms=EB;
CC IsoId=Q8WVK8-2; Sequence=External;
CC Name=7;
CC IsoId=Q03001-6; Sequence=External;
CC Name=8;
CC IsoId=O94833-3; Sequence=VSP_005068, VSP_005069;
CC Name=10;
CC IsoId=O94833-1; Sequence=VSP_005066, VSP_005067;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and
cultured keratinocytes.
CC -!- DISEASE: BPAG1 is an autoantigen of bullous pemphigoid
[IMM:600088], an autoimmune subepithelial skin blistering disease.
CC -!- SIMILARITY: Belongs to the plakophilin or cytokeratin family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 29 spectrin repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF400226; AAL62061.1; -;
CC EMBL; BC016991; AAL16991.1; -;
CC EMBL; AK055189; BAC70870.1; ALT_INIT.
CC EMBL; AK094883; BAC04449.1; ALT_INIT.
CC EMBL; AK096713; BAC04448.1; ALT_INIT.
CC EMBL; AB018271; BAA34448.2; -;
CC EMBL; AL137008; CAC12899.1; -;
CC HGSP; O82040; LK9U.
CC IntAct; O94833; -;
CC Genew; HGNC:1090; DST.
CC MIM; 113810; -;
CC MIM; 600088; -;
CC GO; GO:0005737; Cytoplasm; ISS.
CC GO; GO:0005200; F:structural constituent of cytoskeleton; ISS.
CC GO; GO:0045104; P:intermediate filament cytoskeleton organiza. .; ISS.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR010983; EF_Hand_like.
CC InterPro; IPR003108; GAS2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00036; ehand; 2.
CC Pfam; PF02187; GAS2; 1.
CC Pfam; PF00435; Spectrin; 29.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00054; Eph; 2.
CC SMART; SM00243; GAS2; 1.
CC SMART; SM00150; SPEC; 30.
CC PROSITE; PS00018; EF_HAND; 2.
CC ACTIN-binding; Alternative splicing; Antigen; Calcium;
CC Calcium-binding; Cell adhesion; Coiled coil; Cytoskeleton; Repeat;
CC SH3 domain; Structural protein.
CC REPEAT 264 341 Spectrin 1.
CC REPEAT 349 444 Spectrin 2.
CC REPEAT 450 550 Spectrin 3.
CC DOMAIN 564 616 SH3.
CC REPEAT 935 1031 Spectrin 4.
CC REPEAT 1481 1581 Spectrin 5.
CC REPEAT 1715 1814 Spectrin 6.
CC REPEAT 1930 2008 Spectrin 7.
CC REPEAT 2071 2174 Spectrin 8.

FT REPEAT 2184 2282 Spectrin 9.
FT REPEAT 2294 2382 Spectrin 10.

Query Match 97.9%; Score 190; DB 1; Length 5171;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SNAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 40
DB 2784 SNAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 2822

RESULT 2

Q6ZQ59 PRELIMINARY; PRT; 1589 AA.
ID Q6ZQ59; AC Q6ZQ59;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE MKIAA0728 protein (Fragment).
GN Name=mkIAA0728;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryonic tail;
RX PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT cDNAs identified by screening of terminal sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
DR EMBL; AX129201; BAC98011.1; -;
DR GO; GO:0030056; C:hemidesmosome; IDA.
DR GO; GO:0008090; P:retrograde axon cargo transport; IMP.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00435; Spectrin; 9.
DR SMART; SM00150; SPECT; 12.
FT NON TER 1
SQ SEQUENCE 1589 AA; 180379 MW; CID3CFAA99975454 CRC64;

Query Match 77.8%; Score 151; DB 2; Length 1589;
Best Local Similarity 78.9%; Pred. No. 1.8e-09;
Matches 30; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 MAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 40
DB 491 MAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 528

RESULT 3

BAC98011 PRELIMINARY; PRT; 1589 AA.
ID BAC98011; AC BAC98011;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE MKIAA0728 protein (Fragment).
GN MKIAA0728.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryonic tail;
RX PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of KIAA Gene:

III. The Complete Nucleotide Sequences of 500 Mouse KIAA-homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-fractionated Libraries.";
RL DNA Res. 10:167-180(2003).
DR EMBL; AK129201; BAC98011.1; -;
FT NON TER 1
SQ SEQUENCE 1589 AA; 180379 MW; CID3CFAA99975454 CRC64;

Query Match 77.8%; Score 151; DB 2; Length 1589;
Best Local Similarity 78.9%; Pred. No. 1.8e-09;
Matches 30; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 MAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 40
DB 491 MAPVGRDAETLQKQKTIKAFLLKLEALMASNDNANKT 528

RESULT 4

BPA1_MOUSE STANDARD; PRT; 7389 AA.
ID BPA1_MOUSE; AC Q912U6; Q912U7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Bullous pemphigoid antigen 1, isoforms 1/2/3/4 (BPA) (Hemidesmosomal
DE plaque protein) (Dystonia musculorum protein) (Dystonin).
GN Name=Bst; Synonyms=Bspagi;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
RP STRAIN=BALB/c; TISSUE=Muscle, and Neuron;
RX MEDLINE=21405767; PubMed=11514586; DOI=10.1093/jcb.200012098;
RC Leung C.L., Zheng M., Prater S.M., Liem R.K.H.;
RT "The BPAG1 locus: alternative splicing produces multiple isoforms with
RT distinct cytoskeletal linker domains, including predominant isoforms
RT in neurons and muscles.";
RL J. Cell Biol. 154:691-697(2001).
RN [2]

SEQUENCE OF 6693-7389 FROM N.A. (ISOFORMS 3 AND 4).
RP STRAIN=CS7BL/6J; TISSUE=Fetal skin, and Fetal spinal cord;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

```

CC -|- FUNCTION: Cytoskeletal linker protein. Anchors keratin-containing
CC intermediate filaments to the inner plaque of hemidesmosomes. The
CC proteins may self-aggregate to form filaments or a two-dimensional
CC mesh (By similarity).
CC
CC -|- SUBUNIT: Homodimer. Interacts with the neuronal intermediate
CC filament protein, Prph (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- ALTERNATIVE PRODUCTS:
CC   Event-Alternative splicing; Named isoforms=7;
CC   Name=2; Synonyms=b;
CC   IsoId=Q91ZU6-1; Sequence=Displayed;
CC   Name=1; Synonyms=a;
CC   IsoId=Q91ZU6-2; Sequence=VSP_050483;
CC   Name=3;
CC   IsoId=Q91ZU6-3; Sequence=VSP_050484, VSP_050485, VSP_050486;
CC   Note=No experimental confirmation available;
CC   Name=4;
CC   IsoId=Q91ZU6-4; Sequence=VSP_050485, VSP_050486;
CC   Note=No experimental confirmation available;
CC   Name=5; Synonyms=e;
CC   IsoId=Q91ZU8-1; Sequence=External;
CC   Name=6; Synonyms=n1;
CC   IsoId=Q60824-1; Sequence=External;
CC   Name=7; Synonyms=n2;
CC   IsoId=Q60824-2; Sequence=External;
CC -|- TISSUE SPECIFICITY: Expressed at high levels in the heart and
CC skeletal muscle and at low levels in the skin in the adult.
CC Expressed in the myocardium, skeletal muscle masses, vertebrae
CC cartilage, and epithelia of the tongue of 14.5 day embryos.
CC -|- SIMILARITY: Belongs to the plakin or cytolinker family.
CC -|- SIMILARITY: Contains 1 actin-binding domain.
CC -|- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -|- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -|- SIMILARITY: Contains 9 plectrin repeats.
CC -|- SIMILARITY: Contains 1 SH3 domain.
CC -|- SIMILARITY: Contains 27 spectrin repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF396879; AK83384.1; -.
CC EMBL; AF396878; AK83383.1; -.
CC EMBL; AK051626; BAC34695.1; -.
CC EMBL; AK037206; BAC29753.1; -.
CC HSP; O15149; 1MB8.
CC MGD; MGI:104627; Dst.
CC GO; GO:0005737; Cytoplasm; ISS.
CC GO; GO:0030056; Cytoskeleton; ISS.
CC GO; GO:0005200; F-structural constituent of cytoskeleton; ISS.
CC GO; GO:0045104; P-intermediate filament cytoskeleton organiza. .; ISS.
CC InterPro; IPR001589; Actbind actin.
CC InterPro; IPR001715; Calponin-like.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR0010983; EF-hand.
CC InterPro; IPR003108; GAS2.
CC InterPro; IPR001101; Plectrin_repeat.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00307; CH; 2.
CC Pfam; PF00036; ehfand; 2.
CC Pfam; PF02187; GAS2; 1.
CC Pfam; PF00681; Plectrin; 4.
CC Pfam; PF00435; Spectrin; 26.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00033; CH; 2.
CC SMART; SM00054; EFh; 2.
CC SMART; SM00243; GAS2; 1.
CC SMART; SM00250; PLEC; 9.

```

```

DR SMART; SM00150; SPEC; 32.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00021; CH; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR ACTIN-BINDING; Alternative splicing; Calcium; Calcium-binding;
DR Cell adhesion; Coiled coil; Cytoskeleton; Repeat; SH3 domain;
DR Structural protein.
DR DOMAIN 31 255 Actin-binding.
DR DOMAIN 35 138 CH 1.
DR DOMAIN 151 252 CH 2.
DR REPEAT 590 667 Spectrin 1.
DR REPEAT 675 770 Spectrin 2.
DR REPEAT 889 941 SH3.
DR REPEAT 1260 1356 Spectrin 3.
DR REPEAT 1537 1581 Plectin 1.
DR REPEAT 1582 1619 Plectin 2.
DR REPEAT 1657 1694 Plectin 3.
DR REPEAT 1695 1732 Plectin 4.
DR REPEAT 1735 1770 Plectin 5.
DR REPEAT 1771 1808 Plectin 6.
DR REPEAT 1811 1846 Plectin 7.
DR REPEAT 1847 1884 Plectin 8.
DR REPEAT 1886 1922 Plectin 9.
DR REPEAT 3814 3914 Plectin 4.
DR REPEAT 4053 4152 Spectrin 4.
DR REPEAT 4270 4346 Spectrin 6.
DR REPEAT 4409 4517 Spectrin 7.
DR REPEAT 4522 4620 Spectrin 8.
DR REPEAT 4623 4729 Spectrin 9.
DR REPEAT 4742 4840 Spectrin 10.
DR REPEAT 4851 4949 Spectrin 11.
DR REPEAT 5177 5278 Spectrin 12.
DR REPEAT 5288 5385 Spectrin 13.
DR REPEAT 5397 5497 Spectrin 14.
DR REPEAT 5506 5605 Spectrin 15.
DR REPEAT 5646 5714 Spectrin 16.
DR REPEAT 5725 5824 Spectrin 17.
DR REPEAT 5946 6045 Spectrin 18.
DR REPEAT 6055 6155 Spectrin 19.
DR REPEAT 6165 6265 Spectrin 20.
DR REPEAT 6274 6372 Spectrin 21.
DR REPEAT 6383 6480 Spectrin 22.
DR REPEAT 6492 6592 Spectrin 23.
DR REPEAT 6602 6701 Spectrin 24.
DR REPEAT 6710 6808 Spectrin 25.
DR REPEAT 6826 6914 Spectrin 26.
DR REPEAT 6962 7020 Spectrin 27.
DR REPEAT 7028 7039 EF-hand 1 (Potential).
DR CA_BIND 7064 7075 EF-hand 2 (Potential).
DR DOMAIN 452 486 Coiled coil (Potential).
DR DOMAIN 638 703 Coiled coil (Potential).
DR DOMAIN 737 773 Coiled coil (Potential).
DR DOMAIN 1003 1138 Coiled coil (Potential).
DR DOMAIN 1195 1247 Coiled coil (Potential).
DR DOMAIN 1413 1455 Coiled coil (Potential).
DR DOMAIN 1504 1527 Coiled coil (Potential).
DR DOMAIN 3336 3359 Coiled coil (Potential).
DR DOMAIN 3539 3715 Coiled coil (Potential).
DR DOMAIN 3809 3893 Coiled coil (Potential).
DR DOMAIN 3957 3978 Coiled coil (Potential).
DR DOMAIN 4006 4043 Coiled coil (Potential).
DR DOMAIN 4159 4209 Coiled coil (Potential).
DR DOMAIN 4270 4311 Coiled coil (Potential).
DR DOMAIN 4426 4563 Coiled coil (Potential).
DR DOMAIN 4696 4735 Coiled coil (Potential).
DR DOMAIN 4847 5097 Coiled coil (Potential).
DR DOMAIN 5173 5233 Coiled coil (Potential).
DR DOMAIN 5570 5603 Coiled coil (Potential).
DR DOMAIN 5717 5739 Coiled coil (Potential).
DR DOMAIN 5787 5809 Coiled coil (Potential).
DR DOMAIN 6010 6089 Coiled coil (Potential).
DR DOMAIN 6116 6164 Coiled coil (Potential).
DR DOMAIN 6277 6318 Coiled coil (Potential).

```

```
Query Match      77.8%; Score 151; DB 1; Length 7389;
Best Local Similarity 78.9%; Pred No. 8.7e-09;
Matches 30; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 MAPVGRDATTLOKQKXETIKAFKLUKLEALMASNDNANKT 40
DB 5123 MAPVGRDATTLOKQKACQTKFLKLEALMASNDSANRT 5160

RESULT 5
MACF_MOUSE
ID Q9QXZ0; P97394; STANDARD; PRT; 5327 AA.
AC Q9QXZ0; P97395; P97396;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Microtubule-actin crosslinking factor 1 (Actin cross-linking family
7).
GN Name=Macf1; Synonyms=Macf, Acf7, Acfp7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=BALE/C;
RX MEDLINE=20069791; PubMed=10601340;
RA Leung C.L., Sun D., Zheng M., Knowles D.R., Liem R.K.H.;
RT "Microtubule actin cross-linking factor (MACF): a hybrid of dystonin
RT and dystrophin that can interact with the actin and microtubule
RT cytoskeletons.";
RL J. Cell Biol. 147:1275-1286(1999).
RN [2]
PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC STRAIN=BALE/C; TISSUE=Brain;
RX MEDLINE=97124842; PubMed=8954775;
RA Bernier G., Mathieu M., De Repentigny Y., Vidal S.M., Kothary R.;
RT "Cloning and characterization of mouse ACF7, a novel member of the
RT dysconin subfamily of actin binding proteins.";
RL Genomics 38:19-29(1996).
CC -!- FUNCTION: F-actin-binding protein which may play a role in cross-
CC -!- linking actin to other cytoskeletal proteins. Also binds to
CC microtubules.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC Name=1;
CC IsoId=Q9QXZ0-1; Sequence=Displayed;
CC IsoId=Q9QXZ0-2; Sequence=VSP_000717;
CC Note-Incomplete sequence;
CC Name=3;
CC IsoId=Q9QXZ0-3; Sequence=VSP_000718;
CC TISSUE SPECIFICITY: Expressed mainly in lung, brain, spinal cord,
CC skeletal and cardiac muscle, and skin.
CC -!- SIMILARITY: Contains 1 actin-binding domain.
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 37 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF150755; AAD32244.1; -.
DR EMBL; U67203; AAC52988.1; -.
DR EMBL; U67204; AAC52989.1; -.

```

```
DR EMBL; U67205; AAC52990.1; -.
DR HSP; Q15149; IMB8
DR MGD; MGI1108559; Macf1.
DR GO; GO:0015629; C:actin cytoskeleton; IDA.
DR GO; GO:0015630; C:microtubule cytoskeleton; IDA.
DR GO; GO:0003779; F:actin binding; IDA.
DR GO; GO:0008017; F:microtubule binding; IDA.
DR GO; GO:0006928; P:cell motility; IMP.
DR GO; GO:0007163; P:establishment and/or maintenance of cell po. .; IMP.
DR InterPro; IPR001589; Actbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002048; EF_Hand.
DR InterPro; IPR010983; EF_Hand_like.
DR InterPro; IPR003108; GAS2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF02187; GAS2; 1.
DR ProDom; PD000435; Spectrin; 27.
DR PROSITE; PS000012; EF-hand; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
DR PROSITE; PS00021; CH; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PSS0002; SH3; FALSE_NEG.
KW Actin-binding; Alternative splicing; Calcium-binding; Cytoskeleton;
KW Repeat; SH3 domain.
FT DOMAIN 1 295 Actin-binding (By similarity).
FT DOMAIN 78 181 CH 1.
FT DOMAIN 194 295 CH 2.
FT REPEAT 314 355 Spectrin 1.
FT REPEAT 591 623 Spectrin 2.
FT REPEAT 680 784 Spectrin 3.
FT REPEAT 786 800 Spectrin 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 Spectrin 5.
FT REPEAT 1287 1342 Spectrin 6.
FT REPEAT 1458 1534 Spectrin 7.
FT REPEAT 1593 1660 Spectrin 8.
FT REPEAT 1817 1886 Spectrin 9.
FT REPEAT 1934 2044 Spectrin 10.
FT REPEAT 2262 2282 Spectrin 11.
FT REPEAT 2376 2397 Spectrin 12.
FT REPEAT 2400 2509 Spectrin 13.
FT REPEAT 2557 2617 Spectrin 14.
FT REPEAT 2620 2727 Spectrin 15.
FT REPEAT 2730 2837 Spectrin 16.
FT REPEAT 2840 2944 Spectrin 17.
FT REPEAT 2946 2972 Spectrin 18.
FT REPEAT 2986 3029 Spectrin 19.
FT REPEAT 3086 3162 Spectrin 20.
FT REPEAT 3195 3273 Spectrin 21.
FT REPEAT 3276 3382 Spectrin 22.
FT REPEAT 3385 3491 Spectrin 23.
FT REPEAT 3494 3518 Spectrin 24.
FT REPEAT 3603 3709 Spectrin 25.
FT REPEAT 3722 3817 Spectrin 26.
FT REPEAT 3869 3930 Spectrin 27.
FT REPEAT 3933 4039 Spectrin 28.
FT REPEAT 4042 4149 Spectrin 29.
FT REPEAT 4152 4258 Spectrin 30.
FT REPEAT 4261 4368 Spectrin 31.
FT REPEAT 4371 4477 Spectrin 32.
FT REPEAT 4480 4587 Spectrin 33.
FT REPEAT 4594 4695 Spectrin 34.
FT REPEAT 4698 4804 Spectrin 35.
FT REPEAT 4807 4872 Spectrin 36.
FT REPEAT 4917 4941 Spectrin 37.
FT CA BIND 4987 4998 EF-hand 1 (Potential).
FT CA BIND 5023 5034 EF-hand 2 (Potential).
FT DOMAIN 5173 5180 Poly-Ser.
FT DOMAIN 5252 5267 4 X 4 AA tandem repeats of [GS]-S-R-[AR].

```

FT VARGPLIC 1 73 MSSDEETLSRSCRSRSCRSRSCRSRSCRSRSGSLSPCPPGD
 FT TLPWNLPHLEQKRRKQSDVLDPAERAVRVRA -> EKEFV
 FT QAVEDVLRK (in isoform 1).
 FT /FTid=VSP 000717
 FT MSSDEETLSRSCRSRSCRSRSCRSRSCRSRSGSLSPCPPGD
 FT TLPWNLPHLEQKRRKQSDVLDPAERAVRVRAERDRVQK
 FT TPTKWNKHLKMKRKHINDLYEDLDGHNLI:SLLEVLSGIK
 FT LPRKGRMFRHLQVQIALDQKQKQVKNIRNDITDQ
 FT NKPLTLGLIITLILFO -> MGNSLGCVKPKESIAVPEK
 FT APISPKKVRFRKRWGKILTPASHREBALEGTGVIEET
 FT ETLKLTALPKPEPGVGAHPSPDIFLPGDSAPNSGVGDQ
 FT GMIVQKSFQAEIQTALHLENSSVVGGAWSLSEGMTV
 FT IAHLLNPAERNCEKSVQVLEFFPTACSSRAVLLPQGE
 FT TAVEGGTLLRRHRSSTLPRTDYPSETVDQDPSEGSVG
 FT GRTKGVPAAPPTGSGVIAKSVASSIPKQSGDPIHTEPTHV
 FT LVSCGPIMPASQSLSVSGITVSLPSSGGVSGDGLRLHG
 FT IRPTEPEKSTPPESEEDGTLISLE (in isoform 3).
 FT /FTid=VSP 000718.
 FT L -> P (in Ref. 2; AAC52990).
 FT Q -> H (in Ref. 2).
 FT CONFLICT 393 393
 FT CONFLICT 1882 1882
 FT SEQUENCE 5327 AA; 607972 MW; 5DE4FF5A6514BFFA CRC64;

Query Match 37.6%; Score 73; DB 1; Length 5327;
 Best Local Similarity 46.4%; Pred. No. 23;
 Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 3 MAPVGRDAETLQKQETIKAFKLKLEAL 30

DB 3082 MGAIGRDTSLQSQIEDVRLFLNKIQAL 3109

RESULT 6
 Q89QJ6 PRELIMINARY; PRT; 486 AA.
 ID Q89QJ6
 AC Q89QJ6
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Two-component sensor histidine kinase.
 GN OrderedLocusNames=blr3132;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Teurouka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 Bradyrhizobium japonicum USDA110.";
 RL DNA Res. 9:189-197(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL: AP005946; BAC48397.1; -;
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0005524; F: ATP binding; IEA.
 DR GO: GO:0016301; F: kinase activity; IEA.
 DR GO: GO:0016740; F: transferase activity; IEA.
 DR GO: GO:0000155; F: two-component sensor molecule activity; IEA.
 DR GO: GO:0007600; P: sensory perception; IEA.
 DR GO: GO:0007165; P: signal transduction; IEA.
 DR InterPro: IPR003594; ATPbind ATPase.
 DR InterPro: IPR004356; Bact_sens_pr_C.
 DR InterPro: IPR003660; HAMF.
 DR InterPro: IPR005467; His_kinase.
 DR InterPro: IPR003661; His_kin_N.
 DR InterPro: IPR009082; His_kin_homodim.
 DR Pfam: PF00672; HAMF; 1.
 DR Pfam: PF02518; HATPase_c; 1.

DR Pfam: PF00512; HiskA; 1.
 DR PRINTS: PRO0344; ECTRLENSOR.
 DR SMART: SM00304; HAMF; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00388; HiskA; 1.
 DR PROSITE: PS00885; HAMF; 1.
 DR PROSITE: PS01019; HIS_KIN; 1.
 KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
 KW Transferase; Transmembrane.
 SQ SEQUENCE 486 AA; 52550 MW; 329B8CC0C94D017C CRC64;

Query Match 35.6%; Score 69; DB 2; Length 486;
 Best Local Similarity 41.7%; Pred. No. 6.3;
 Matches 15; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

QY 5 FVGRDAETLQKQETIKAFKLKLEALMAS---NDN 36

DB 209 FVGRSGDLDRLAENLNAMLERIEALMAGLKEVSDN 244

RESULT 7
 BAA32310 PRELIMINARY; PRT; 4433 AA.
 ID BAA32310
 AC BAA32310;
 DT 02-MAR-2004 (T-EMBLrel. 27, Created)
 DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
 DE KIAA0465 protein (Fragment).
 GN KIAA0465.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98116662; PubMed=9455484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
 RA Nomura N., Ohara O.;
 RT "Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain.";
 RL DNA Res. 4:345-349(1997).
 DR EMBL: AB007934; BAA32310.3; -;
 FT NON TER
 FT SEQUENCE 4433 AA; 505640 MW; 329C01910D97EB6C CRC64;

Query Match 34.5%; Score 67; DB 2; Length 4433;
 Best Local Similarity 42.9%; Pred. No. 1e+02;
 Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 3 MAPVGRDAETLQKQETIKAFKLKLEAL 30

DB 2195 MGAIGRDTSLQSQIEDVRLFLNKIHVL 2222

RESULT 8
 MACF_HUMAN STANDARD; PRT; 5430 AA.
 ID MACF_HUMAN
 AC Q9UPN3; Q75053; Q8WXY2; Q9H540; Q9UKP0; Q9ULG9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Microtubule-actin crosslinking factor 1, isoforms 1/2/3 (actin cross-
 DE linking family protein 7) (Macrophilin 1) (trabeculin-alpha) (620 kDa
 DE actin-binding protein) (ABP620).
 GN Name=MACF1; Synonyms=ACF7, ABP620, KIAA0465, KIAA1251;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=20001959; PubMed=10529403; DOI=10.1006/bbrc.1999.1538;

RA Okuda T., Matsuda S., Nakatsugawa S., Ichigotani Y., Iwahashi N.,
RA Takahashi M., Ishigaki T., Hamaguchi M.;
RT "Molecular cloning of macrophin, a human homologue of Drosophila
RT kaxapo with a close structural similarity to plectin and dystrophin.";
RL Biochem. Biophys. Res. Commun. 264:568-574(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
EX MEDLINE=20026884; PubMed=10559237;
RA Sun Y., Zhang J., Kraeft S.-K., Auclair D., Chang M.-S., Liu Y.,
RA Sutherland R., Salgia R., Griffin J.D., Ferland L.H., Chen L.B.;
RT "Molecular cloning and characterization of human trabeculin-alpha, a
RT giant protein defining a new family of actin-binding proteins.";
RN J. Biol. Chem. 274:33522-33530(1999).
RN [3]
RP SEQUENCE OF 182-4812 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=21833812; PubMed=11845288;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "MACF1 gene structure: a hybrid of plectin and dystrophin.";
RN Mamm. Genome 12:852-861(2001).
RN [4]
RP SEQUENCE OF 668-2350 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [5]
RP SEQUENCE OF 1544-5057 FROM N.A.
RA Corby N.;
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 3312-5430 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RN DNA Res. 9:99-106(2002).
RN [7]
RP SEQUENCE OF 3734-5430 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
RA Nakajima D., Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
RT from human brain.";
RN DNA Res. 4:345-349(1997).
RL -!- FUNCTION: F-actin-binding protein which may play a role in cross-
CC linking actin to other cytoskeletal proteins. Also binds to
CC microtubules (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=VSP_007341;
CC Name=3;
CC IsoId=Q9UPN3-3; Sequence=Not described;
CC Name=4;
CC IsoId=Q96PK2-1; Sequence=External;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the plectin or cytolinker family.
CC -!- SIMILARITY: Contains 1 actin-binding domain.
CC -!- SIMILARITY: Contains 2 actin-binding domain.
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 37 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB029290; BAA83821.1; -
CC EMBL; AF141968; AAP06360.1; -
CC EMBL; AF325341; AAL39000.1; -
CC EMBL; AF325330; AAL39000.1; JOINED.
CC EMBL; AF325331; AAL39000.1; JOINED.
CC EMBL; AF325332; AAL39000.1; JOINED.
CC EMBL; AF325333; AAL39000.1; JOINED.
CC EMBL; AF325334; AAL39000.1; JOINED.
CC EMBL; AF325335; AAL39000.1; JOINED.
CC EMBL; AF325336; AAL39000.1; JOINED.
CC EMBL; AF325339; AAL39000.1; JOINED.
CC EMBL; AF325340; AAL39000.1; JOINED.
CC EMBL; AB033077; BAA86565.1; -
CC EMBL; AL137853; CAC13920.1; -
CC EMBL; AB007934; BAA32310.2; -
CC PIR; T00079; T00079.
CC HSP; Q15149; LMB8.
CC Genew; HGNC:13664; MACF1.
CC MTM; 608271; -
CC GO; GO:0005856; Cytoskeleton; NAS.
CC GO; GO:0003780; Fractin cross-linking activity; NAS.
CC GO; GO:0005509; F-actin ion binding; NAS.
CC GO; GO:0008017; F-actin binding; NAS.
CC InterPro; IPR001589; Actbind actinin.
CC InterPro; IPR001715; Calponin-like.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR010983; EF-Hand-like.
CC InterPro; IPR003108; GAS2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00307; CH; 2.
CC Pfam; PF00036; ehand; 2.
CC Pfam; PF02187; GAS2; 1.
CC Pfam; PF00435; Spectrin; 27.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00033; CH; 2.
CC SMART; SM00054; EFH; 2.
CC SMART; SM00243; GAS2; 1.
CC SMART; SM00150; SPEC; 33.
CC PROSITE; PS00019; ACTININ 1; 1.
CC PROSITE; PS00020; ACTININ 2; FALSE_NEG.
CC PROSITE; PS00021; CH; 2.
CC PROSITE; PS00018; EF HAND; 2.
CC PROSITE; PS00002; SH3; FALSE_NEG.
CC Actin-binding; Alternative splicing; Calcium; Calcium-binding;
CC Coiled coil; Cytoskeleton; Repeat; SH3 domain.
CC Coiled coil; Actin-binding.
CC DOMAIN 1 295
CC FT DOMAIN 78 181
CC FT DOMAIN 194 295
CC FT DOMAIN 243 265
CC FT DOMAIN 477 529
CC FT DOMAIN 722 751
CC FT DOMAIN 816 843
CC FT DOMAIN 1013 1118
CC FT DOMAIN 1164 1191
CC FT DOMAIN 1399 1690
CC FT DOMAIN 1780 1843
CC FT DOMAIN 1975 2005
CC FT DOMAIN 2039 2312
CC FT DOMAIN 2385 2417
CC FT DOMAIN 2544 2695
CC FT DOMAIN 2760 2838
CC FT DOMAIN 2911 3001
CC FT DOMAIN 3130 3164
CC FT DOMAIN 3244 3277
CC FT DOMAIN 3418 3492

FT	DOMAIN	3596	3666	Coiled coil (Potential).	CC	Name=4;
FT	DOMAIN	3786	3806	Coiled coil (Potential).	CC	IsoId=Q96PK2-1; Sequence=Displayed;
FT	DOMAIN	3852	3931	Coiled coil (Potential).	CC	Name=1;
FT	DOMAIN	3967	3987	Coiled coil (Potential).	CC	IsoId=Q9UPN3-1; Sequence=External;
FT	DOMAIN	4084	4218	Coiled coil (Potential).	CC	Name=2;
FT	DOMAIN	4343	4378	Coiled coil (Potential).	CC	IsoId=Q9UPN3-2; Sequence=External;
FT	DOMAIN	4408	4437	Coiled coil (Potential).	CC	Name=3;
FT	DOMAIN	4468	4498	Coiled coil (Potential).	CC	IsoId=Q9UPN3-3; Sequence=External;
FT	DOMAIN	4907	4935	Coiled coil (Potential).	CC	TISSUE SPECIFICITY: Expressed in heart, lung, pituitary and placenta, not found in brain, kidney, liver, pancreas or skeletal muscle.
FT	DOMAIN	5044	5067	Coiled coil (Potential).	CC	-1- SIMILARITY: Belongs to the plakin or cytolinker family.
FT	REPEAT	314	355	Spectrin 1.	CC	-1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
FT	REPEAT	591	623	Spectrin 2.	CC	-1- SIMILARITY: Contains 19 plectrin repeats.
FT	REPEAT	680	784	Spectrin 3.	CC	-1- SIMILARITY: Contains 32 spectrin repeats.
FT	REPEAT	786	800	Spectrin 4.	CC	-----
FT	DOMAIN	871	923	SH3.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
FT	REPEAT	1250	1272	Spectrin 5.	CC	-----
FT	REPEAT	1287	1342	Spectrin 6.	CC	EMBL; AF317696; AAL09459.1; --
FT	REPEAT	1455	1534	Spectrin 7.	CC	EMBL; AF325341; AAL38997.1; JOINED.
FT	REPEAT	1547	1659	Spectrin 8.	CC	EMBL; AF325333; AAL38997.1; JOINED.
FT	REPEAT	1815	1891	Spectrin 9.	CC	EMBL; AF325335; AAL38997.1; JOINED.
FT	REPEAT	1932	2042	Spectrin 10.	CC	EMBL; AF325336; AAL38997.1; JOINED.
FT	REPEAT	2260	2280	Spectrin 11.	CC	EMBL; AF325339; AAL38997.1; JOINED.
FT	REPEAT	2372	2395	Spectrin 12.	CC	EMBL; AF325340; AAL38997.1; JOINED.
FT	REPEAT	2398	2507	Spectrin 13.	CC	HSSP; P15924; 1LM5.
FT	REPEAT	2510	2618	Spectrin 14.	CC	Genew; HGNC:13664; MACF1.
FT	REPEAT	2621	2728	Spectrin 15.	CC	MIM; 608271; --
FT	REPEAT	2731	2838	Spectrin 16.	CC	GO; GO:0005856; C:cytoskeleton; ISS.
FT	REPEAT	2841	2945	Spectrin 17.	CC	GO; GO:0005509; F:calcium ion binding; ISS.
FT	REPEAT	2987	3024	Spectrin 18.	CC	GO; GO:008017; F:microtubule binding; ISS.
FT	REPEAT	3136	3163	Spectrin 19.	CC	InterPro; IPR002048; EF-hand.
FT	REPEAT	3187	3274	Spectrin 20.	CC	InterPro; IPR010983; EF-hand_like.
FT	REPEAT	3277	3383	Spectrin 21.	CC	InterPro; IPR003108; GAS2.
FT	REPEAT	3386	3492	Spectrin 22.	CC	InterPro; IPR001101; Plectrin_repeat.
FT	REPEAT	3495	3601	Spectrin 23.	CC	InterPro; IPR002017; Spectrin.
FT	REPEAT	3604	3673	Spectrin 24.	CC	Pfam; PF00036; ehand; 2.
FT	REPEAT	3673	3746	Spectrin 25.	CC	Pfam; PF02187; GAS2; 1.
FT	REPEAT	3713	3819	Spectrin 26.	CC	Pfam; PF00681; Plectrin; 11.
FT	REPEAT	3832	3927	Spectrin 27.	CC	Pfam; PF00435; Spectrin; 26.
FT	REPEAT	3982	4043	Spectrin 27.	CC	ProDom; PD000012; EF-hand; 1.
Query Match 34.5%; Score 67; DB 1; Length 5430;						
Best Local Similarity 42.9%; Pred. No. 1.3e+02;						
Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;						
QY	3 MAPVGRDAETLQKQETIKAFKKLEAL 30					
DB	3083 MGAIGRDTDSLSQIEDVLFNKLHVL 3110					
RESULT 9						
MAC4	HUMAN					
ID	MAC4_HUMAN					
AC	Q96PK2; Q8WXY1; STANDARD; PRT; 5938 AA.					
DT	10-OCT-2003 (Rel. 42, Created)					
DT	10-OCT-2003 (Rel. 42, Last sequence update)					
DT	01-OCT-2004 (Rel. 45, Last annotation update)					
DE	Microtubule-actin crosslinking factor 1, isoform 4.					
GN	Name=MACF1; Synonyms=ACF7, ABP620, KIAA0465, KIAA1251;					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.					
RX	MEDLINE-21833812; PubMed-11845288;					
RA	Gong T.-W.L., Besirli C.G., Lomax M.I.;					
RT	"MACF1 gene structure: a hybrid of plectin and dystrophin.";					
RL	Mamm. Genome 12:852-861(2001).					
CC	-1- FUNCTION: May play a role in cross-linking cytoskeletal proteins					
CC	by binding intermediate filaments to the N-terminal plectin					
CC	repeats and microtubules to the C-terminus.					
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).					
CC	-1- ALTERNATIVE PRODUCTS:					
CC	Event=Alternative splicing; Named isoforms=4;					

FT DOMAIN 4845 4880 Coiled coil (Potential).
 FT DOMAIN 4910 4939 Coiled coil (Potential).
 FT DOMAIN 4970 5000 Coiled coil (Potential).
 FT DOMAIN 5437 5469 Coiled coil (Potential).
 FT DOMAIN 5546 5569 Coiled coil (Potential).
 FT CA_BIND 5598 5610 EF-hand 1 (Potential).
 FT CA_BIND 5634 5646 EF-hand 2 (Potential).
 FT REPEAT 12 49 Plectin 1.
 FT REPEAT 53 88 Plectin 2.
 FT REPEAT 89 126 Plectin 3.
 FT REPEAT 130 164 Plectin 4.
 FT REPEAT 166 202 Plectin 5.
 FT REPEAT 203 240 Plectin 6.
 FT REPEAT 243 278 Plectin 7.
 FT REPEAT 279 316 Plectin 8.
 FT REPEAT 318 354 Plectin 9.
 FT REPEAT 375 412 Plectin 10.
 FT REPEAT 414 451 Plectin 11.
 FT REPEAT 453 490 Plectin 12.
 FT REPEAT 492 529 Plectin 13.
 FT REPEAT 531 568 Plectin 14.
 FT REPEAT 570 607 Plectin 15.
 FT REPEAT 609 646 Plectin 16.
 FT REPEAT 648 685 Plectin 17.
 FT REPEAT 687 724 Plectin 18.
 FT REPEAT 726 763 Plectin 19.
 FT REPEAT 765 802 Plectin 20.
 FT REPEAT 804 841 Plectin 21.
 FT REPEAT 843 880 Plectin 22.
 FT REPEAT 882 919 Plectin 23.
 FT REPEAT 921 958 Plectin 24.
 FT REPEAT 960 997 Plectin 25.
 FT REPEAT 999 1036 Plectin 26.
 FT REPEAT 1038 1075 Plectin 27.
 FT REPEAT 1077 1114 Plectin 28.
 FT REPEAT 1116 1153 Plectin 29.
 FT REPEAT 1155 1192 Plectin 30.
 FT REPEAT 1194 1231 Plectin 31.
 FT REPEAT 1233 1270 Plectin 32.
 FT REPEAT 1272 1309 Plectin 33.
 FT REPEAT 1311 1348 Plectin 34.
 FT REPEAT 1350 1387 Plectin 35.
 FT REPEAT 1389 1426 Plectin 36.
 FT REPEAT 1428 1465 Plectin 37.
 FT REPEAT 1467 1504 Plectin 38.
 FT REPEAT 1506 1543 Plectin 39.
 FT REPEAT 1545 1582 Plectin 40.
 FT REPEAT 1584 1621 Plectin 41.
 FT REPEAT 1623 1660 Plectin 42.
 FT REPEAT 1662 1699 Plectin 43.
 FT REPEAT 1701 1738 Plectin 44.
 FT REPEAT 1740 1777 Plectin 45.
 FT REPEAT 1779 1816 Plectin 46.
 FT REPEAT 1818 1855 Plectin 47.
 FT REPEAT 1857 1894 Plectin 48.
 FT REPEAT 1896 1933 Plectin 49.
 FT REPEAT 1935 1972 Plectin 50.
 FT REPEAT 1974 2011 Plectin 51.
 FT REPEAT 2013 2050 Plectin 52.
 FT REPEAT 2052 2089 Plectin 53.
 FT REPEAT 2091 2128 Plectin 54.
 FT REPEAT 2130 2167 Plectin 55.
 FT REPEAT 2169 2206 Plectin 56.
 FT REPEAT 2208 2245 Plectin 57.
 FT REPEAT 2247 2284 Plectin 58.
 FT REPEAT 2286 2323 Plectin 59.
 FT REPEAT 2325 2362 Plectin 60.
 FT REPEAT 2364 2401 Plectin 61.
 FT REPEAT 2403 2440 Plectin 62.
 FT REPEAT 2442 2479 Plectin 63.
 FT REPEAT 2481 2518 Plectin 64.
 FT REPEAT 2520 2557 Plectin 65.
 FT REPEAT 2559 2596 Plectin 66.
 FT REPEAT 2598 2635 Plectin 67.
 FT REPEAT 2637 2674 Plectin 68.
 FT REPEAT 2676 2713 Plectin 69.
 FT REPEAT 2715 2752 Plectin 70.
 FT REPEAT 2754 2791 Plectin 71.
 FT REPEAT 2793 2830 Plectin 72.
 FT REPEAT 2832 2869 Plectin 73.
 FT REPEAT 2871 2908 Plectin 74.
 FT REPEAT 2910 2947 Plectin 75.
 FT REPEAT 2949 2986 Plectin 76.
 FT REPEAT 2988 3025 Plectin 77.
 FT REPEAT 3027 3064 Plectin 78.
 FT REPEAT 3066 3103 Plectin 79.
 FT REPEAT 3105 3142 Plectin 80.
 FT REPEAT 3144 3181 Plectin 81.
 FT REPEAT 3183 3220 Plectin 82.
 FT REPEAT 3222 3259 Plectin 83.
 FT REPEAT 3261 3298 Plectin 84.
 FT REPEAT 3300 3337 Plectin 85.
 FT REPEAT 3339 3376 Plectin 86.
 FT REPEAT 3378 3415 Plectin 87.
 FT REPEAT 3417 3454 Plectin 88.
 FT REPEAT 3456 3493 Plectin 89.
 FT REPEAT 3495 3532 Plectin 90.
 FT REPEAT 3534 3571 Plectin 91.
 FT REPEAT 3573 3610 Plectin 92.
 FT REPEAT 3612 3649 Plectin 93.
 FT REPEAT 3651 3688 Plectin 94.
 FT REPEAT 3690 3727 Plectin 95.
 FT REPEAT 3729 3766 Plectin 96.
 FT REPEAT 3768 3805 Plectin 97.
 FT REPEAT 3807 3844 Plectin 98.
 FT REPEAT 3846 3883 Plectin 99.
 FT REPEAT 3885 3922 Plectin 100.
 FT REPEAT 3924 3961 Plectin 101.
 FT REPEAT 3963 4000 Plectin 102.
 FT REPEAT 4002 4039 Plectin 103.
 FT REPEAT 4041 4078 Plectin 104.
 FT REPEAT 4080 4117 Plectin 105.
 FT REPEAT 4119 4156 Plectin 106.
 FT REPEAT 4158 4195 Plectin 107.
 FT REPEAT 4197 4234 Plectin 108.
 FT REPEAT 4236 4273 Plectin 109.
 FT REPEAT 4275 4312 Plectin 110.
 FT REPEAT 4314 4351 Plectin 111.
 FT REPEAT 4353 4390 Plectin 112.
 FT REPEAT 4392 4429 Plectin 113.
 FT REPEAT 4431 4468 Plectin 114.
 FT REPEAT 4470 4507 Plectin 115.
 FT REPEAT 4509 4546 Plectin 116.
 FT REPEAT 4548 4585 Plectin 117.
 FT REPEAT 4587 4624 Plectin 118.
 FT REPEAT 4626 4663 Plectin 119.
 FT REPEAT 4665 4702 Plectin 120.
 FT REPEAT 4704 4741 Plectin 121.
 FT REPEAT 4743 4780 Plectin 122.
 FT REPEAT 4782 4819 Plectin 123.
 FT REPEAT 4821 4858 Plectin 124.
 FT REPEAT 4860 4897 Plectin 125.
 FT REPEAT 4899 4936 Plectin 126.
 FT REPEAT 4938 4975 Plectin 127.
 FT REPEAT 4977 5014 Plectin 128.
 FT REPEAT 5016 5053 Plectin 129.
 FT REPEAT 5055 5092 Plectin 130.
 FT REPEAT 5094 5131 Plectin 131.
 FT REPEAT 5133 5170 Plectin 132.
 FT REPEAT 5172 5209 Plectin 133.
 FT REPEAT 5211 5248 Plectin 134.
 FT REPEAT 5250 5287 Plectin 135.
 FT REPEAT 5289 5326 Plectin 136.
 FT REPEAT 5328 5365 Plectin 137.
 FT REPEAT 5367 5404 Plectin 138.
 FT REPEAT 5406 5443 Plectin 139.
 FT REPEAT 5445 5482 Plectin 140.
 FT REPEAT 5484 5521 Plectin 141.
 FT REPEAT 5523 5560 Plectin 142.
 FT REPEAT 5562 5599 Plectin 143.
 FT REPEAT 5601 5638 Plectin 144.
 FT REPEAT 5640 5677 Plectin 145.
 FT REPEAT 5679 5716 Plectin 146.
 FT REPEAT 5718 5755 Plectin 147.
 FT REPEAT 5757 5794 Plectin 148.
 FT REPEAT 5796 5833 Plectin 149.
 FT REPEAT 5835 5872 Plectin 150.
 FT REPEAT 5874 5911 Plectin 151.
 FT REPEAT 5913 5950 Plectin 152.
 FT REPEAT 5952 5989 Plectin 153.
 FT REPEAT 5991 6028 Plectin 154.
 FT REPEAT 6030 6067 Plectin 155.
 FT REPEAT 6069 6106 Plectin 156.
 FT REPEAT 6108 6145 Plectin 157.
 FT REPEAT 6147 6184 Plectin 158.
 FT REPEAT 6186 6223 Plectin 159.
 FT REPEAT 6225 6262 Plectin 160.
 FT REPEAT 6264 6301 Plectin 161.
 FT REPEAT 6303 6340 Plectin 162.
 FT REPEAT 6342 6379 Plectin 163.
 FT REPEAT 6381 6418 Plectin 164.
 FT REPEAT 6420 6457 Plectin 165.
 FT REPEAT 6459 6496 Plectin 166.
 FT REPEAT 6498 6535 Plectin 167.
 FT REPEAT 6537 6574 Plectin 168.
 FT REPEAT 6576 6613 Plectin 169.
 FT REPEAT 6615 6652 Plectin 170.
 FT REPEAT 6654 6691 Plectin 171.
 FT REPEAT 6693 6730 Plectin 172.
 FT REPEAT 6732 6769 Plectin 173.
 FT REPEAT 6771 6808 Plectin 174.
 FT REPEAT 6810 6847 Plectin 175.
 FT REPEAT 6849 6886 Plectin 176.
 FT REPEAT 6888 6925 Plectin 177.
 FT REPEAT 6927 6964 Plectin 178.
 FT REPEAT 6966 7003 Plectin 179.
 FT REPEAT 7005 7042 Plectin 180.
 FT REPEAT 7044 7081 Plectin 181.
 FT REPEAT 7083 7120 Plectin 182.
 FT REPEAT 7122 7159 Plectin 183.
 FT REPEAT 7161 7198 Plectin 184.
 FT REPEAT 7200 7237 Plectin 185.
 FT REPEAT 7239 7276 Plectin 186.
 FT REPEAT 7278 7315 Plectin 187.
 FT REPEAT 7317 7354 Plectin 188.
 FT REPEAT 7356 7393 Plectin 189.
 FT REPEAT 7395 7432 Plectin 190.
 FT REPEAT 7434 7471 Plectin 191.
 FT REPEAT 7473 7510 Plectin 192.
 FT REPEAT 7512 7549 Plectin 193.
 FT REPEAT 7551 7588 Plectin 194.
 FT REPEAT 7590 7627 Plectin 195.
 FT REPEAT 7629 7666 Plectin 196.
 FT REPEAT 7668 7705 Plectin 197.
 FT REPEAT 7707 7744 Plectin 198.
 FT REPEAT 7746 7783 Plectin 199.
 FT REPEAT 7785 7822 Plectin 200.
 FT REPEAT 7824 7861 Plectin 201.
 FT REPEAT 7863 7900 Plectin 202.
 FT REPEAT 7902 7939 Plectin 203.
 FT REPEAT 7941 7978 Plectin 204.
 FT REPEAT 7980 8017 Plectin 205.
 FT REPEAT 8019 8056 Plectin 206.
 FT REPEAT 8058 8095 Plectin 207.
 FT REPEAT 8097 8134 Plectin 208.
 FT REPEAT 8136 8173 Plectin 209.
 FT REPEAT 8175 8212 Plectin 210.
 FT REPEAT 8214 8251 Plectin 211.
 FT REPEAT 8253 8290 Plectin 212.
 FT REPEAT 8292 8329 Plectin 213.
 FT REPEAT 8331 8368 Plectin 214.
 FT REPEAT 8370 8407 Plectin 215.
 FT REPEAT 8409 8446 Plectin 216.
 FT REPEAT 8448 8485 Plectin 217.
 FT REPEAT 8487 8524 Plectin 218.
 FT REPEAT 8526 8563 Plectin 219.
 FT REPEAT 8565 8602 Plectin 220.
 FT REPEAT 8604 8641 Plectin 221.
 FT REPEAT 8643 8680 Plectin 222.
 FT REPEAT 8682 8719 Plectin 223.
 FT REPEAT 8721 8758 Plectin 224.
 FT REPEAT 8760 8797 Plectin 225.
 FT REPEAT 8799 8836 Plectin 226.
 FT REPEAT 8838 8875 Plectin 227.
 FT REPEAT 8877 8914 Plectin 228.
 FT REPEAT 8916 8953 Plectin 229.
 FT REPEAT 8955 8992 Plectin 230.
 FT REPEAT 8994 9031 Plectin 231.
 FT REPEAT 9033 9070 Plectin 232.
 FT REPEAT 9072 9109 Plectin 233.
 FT REPEAT 9111 9148 Plectin 234.
 FT REPEAT 9150 9187 Plectin 235.
 FT REPEAT 9189 9226 Plectin 236.
 FT REPEAT 9228 9265 Plectin 237.
 FT REPEAT 9267 9304 Plectin 238.
 FT REPEAT 9306 9343 Plectin 239.
 FT REPEAT 9345 9382 Plectin 240.
 FT REPEAT 9384 9421 Plectin 241.
 FT REPEAT 9423 9460 Plectin 242.
 FT REPEAT 9462 9499 Plectin 243.
 FT REPEAT 9501 9538 Plectin 244.
 FT REPEAT 9540 9577 Plectin 245.
 FT REPEAT 9579 9616 Plectin 246.
 FT REPEAT 9618 9655 Plectin 247.
 FT REPEAT 9657 9694 Plectin 248.
 FT REPEAT 9696 9733 Plectin 249.
 FT REPEAT 9735 9772 Plectin 250.
 FT REPEAT 9774 9811 Plectin 251.
 FT REPEAT 9813 9850 Plectin 252.
 FT REPEAT 9852 9889 Plectin 253.
 FT REPEAT 9891 9928 Plectin 254.
 FT REPEAT 9930 9967 Plectin 255.
 FT REPEAT 9969 10006 Plectin 256.
 FT REPEAT 10008 10045 Plectin 257.
 FT REPEAT 10047 10084 Plectin 258.
 FT REPEAT 10086 10123 Plectin 259.
 FT REPEAT 10125 10162 Plectin 260.
 FT REPEAT 10164 10201 Plectin 261.
 FT REPEAT 10203 10240 Plectin 262.
 FT REPEAT 10242 10279 Plectin 263.
 FT REPEAT 10281 10318 Plectin 264.
 FT REPEAT 10320 10357 Plectin 265.
 FT REPEAT 10359 10396 Plectin 266.
 FT REPEAT 10398 10435 Plectin 267.
 FT REPEAT 10437 10474 Plectin 268.
 FT REPEAT 10476 10513 Plectin 269.
 FT REPEAT 10515 10552 Plectin 270.
 FT REPEAT 10554 10591 Plectin 271.
 FT REPEAT 10593 10630 Plectin 272.
 FT REPEAT 10632 10669 Plectin 273.
 FT REPEAT 10671 10708 Plectin 274.
 FT REPEAT 10710 10747 Plectin 275.
 FT REPEAT 10749 10786 Plectin 276.
 FT REPEAT 10788 10825 Plectin 277.
 FT REPEAT 10827 10864 Plectin 278.
 FT REPEAT 10866 10903 Plectin 279.
 FT REPEAT 10905 10942 Plectin 280.
 FT REPEAT 10944 10981 Plectin 281.
 FT REPEAT 10983 11020 Plectin 282.
 FT REPEAT 11022 11059 Plectin 283.
 FT REPEAT 11061 11098 Plectin 284.
 FT REPEAT 11100 11137 Plectin 285.
 FT REPEAT 11139 11176 Plectin 286.
 FT REPEAT 11178 11215 Plectin 287.
 FT REPEAT 11217 11254 Plectin 288.
 FT REPEAT 11256 11293 Plectin 289.
 FT REPEAT 11295 11332 Plectin 290.
 FT REPEAT 11334 11371 Plectin 291.
 FT REPEAT 11373 11410 Plectin 292.
 FT REPEAT 11412 11449 Plectin 293.
 FT REPEAT 11451 11488 Plectin 294.
 FT REPEAT 11490 11527 Plectin 295.
 FT REPEAT 11529 11566 Plectin 296.
 FT REPEAT 11568 11605 Plectin 297.
 FT REPEAT 11607 11644 Plectin 298.
 FT REPEAT 11646 11683 Plectin 299.
 FT REPEAT 11685 11722 Plectin 300.
 FT REPEAT 11724 11761 Plectin 301.
 FT REPEAT 11763 11800 Plectin 302.
 FT REPEAT 11802 11839 Plectin 303.
 FT REPEAT 11841 11878 Plectin 304.
 FT REPEAT 11880 11917 Plectin 305.
 FT REPEAT 11919 11956 Plectin 306.
 FT REPEAT 11958 12000 Plectin 307.
 FT REPEAT 12002 12045 Plectin 308.
 FT REPEAT 12047 12090 Plectin 309.
 FT REPEAT 12092 12135 Plectin 310.
 FT REPEAT 12137 12180 Plectin 311.
 FT REPEAT 12182 12225 Plectin 312.
 FT REPEAT 12227 12270 Plectin 313.
 FT REPEAT 12272 12315 Plectin 314.
 FT REPEAT 12317 12360 Plectin 315.
 FT REPEAT 12362 12405 Plectin 316.
 FT REPEAT 12407 12450 Plectin 317.
 FT REPEAT 12452 12495 Plectin 318.
 FT REPEAT 12497 12540 Plectin 319.
 FT REPEAT 12542 12585 Plectin 320.
 FT REPEAT 12587 12630 Plectin 321.
 FT REPEAT 12632 12675 Plectin 322.
 FT REPEAT 12677 12720 Plectin 323.
 FT REPEAT 12722 12765 Plectin 324.
 FT REPEAT 12767 12810 Plectin 325.
 FT REPEAT 12812 12855 Plectin 326.
 FT REPEAT 12857 12900 Plectin 327.
 FT REPEAT 12902 12945 Plectin 328.
 FT REPEAT 12947 12990 Plectin 329.
 FT REPEAT 12992 13035 Plectin 330.
 FT REPEAT 13037 13080 Plectin 331.
 FT REPEAT 13082 13125 Plectin 332.
 FT REPEAT 13127 13170 Plectin 333.
 FT REPEAT 13172 13215 Plectin 334.
 FT REPEAT 13217 13260 Plectin 335.
 FT REPEAT 13262 13305 Plectin 336.
 FT REPEAT 13307 13350 Plectin 337.
 FT REPEAT 13352 13395 Plectin 338.
 FT REPEAT 13397 13440 Plectin 339.
 FT REPEAT 13442 13485 Plectin 340.
 FT REPEAT 13487 13530 Plectin 341.
 FT REPEAT 13532 13575 Plectin 342.
 FT REPEAT 13577 13620 Plectin 343.
 FT REPEAT 13622 13665 Plectin 344.
 FT REPEAT 13667 13710 Plectin 345.
 FT REPEAT 13712 13755 Plectin 346.
 FT REPEAT 13757 13800 Plectin 347.
 FT REPEAT 13802 13845 Plectin 348.
 FT REPEAT 13847 13890 Plectin 349.
 FT REPEAT 13892 13935 Plectin 350.
 FT REPEAT 13937 13980 Plectin 351.
 FT REPEAT 13982 14025 Plectin 352.
 FT REPEAT 14027 14070 Plectin 353.
 FT REPEAT 14072 14115 Plectin 354.
 FT REPEAT 14117 14160 Plectin 355.
 FT REPEAT 14162 14205 Plectin 356.
 FT REPEAT 14207 14250 Plectin 357.
 FT REPEAT 14252 14295 Plectin 358.
 FT REPEAT 14297 14340 Plectin 359.
 FT REPEAT 14342 14385 Plectin 360.
 FT REPEAT 14387 14430 Plectin 361.
 FT REPEAT 14432 14475 Plectin 362.
 FT REPEAT 14477 14520 Plectin 363.
 FT REPEAT 14522 14565 Plectin 364.
 FT REPEAT 14567 14610 Plectin 365.
 FT REPEAT 14612 14655 Plectin 366.
 FT REPEAT 14657 14700 Plectin 367.
 FT REPEAT 14702 14745 Plectin 368.
 FT REPEAT 14747 14790 Plectin 369.
 FT REPEAT 14792 14835 Plectin 370.
 FT REPEAT 14837 14880 Plectin 371.
 FT REPEAT 14882 14925 Plectin 372.
 FT REPEAT 14927 14970 Plectin 373.
 FT REPEAT 14972 15015 Plectin 374.
 FT REPEAT 15017 15060 Plectin 375.
 FT REPEAT 15062 15105 Plectin 376.
 FT REPEAT 15107 15150 Plectin 377.
 FT REPEAT 15152 15195 Plectin 378.
 FT REPEAT 15197 15240 Plectin 379.
 FT REPEAT 15242 15285 Plectin 380.
 FT REPEAT 15287 15330 Plectin 381.
 FT REPEAT 15332 15375 Plectin 382.
 FT REPEAT 15377 15420 Plectin 383.
 FT REPEAT 15422 15465 Plectin 384.
 FT REPEAT 15467 15510 Plectin 385.
 FT REPEAT 15512 15555 Plectin 386.
 FT REPEAT 15557 15600 Plectin 387.
 FT REPEAT 15602 15645 Plectin 388.
 FT REPEAT 15647 15690 Plectin 389.
 FT REPEAT 15692 15735 Plectin 390.
 FT REPEAT 15737 15780 Plectin 391.
 FT REPEAT 15782 15825 Plectin 392.
 FT REPEAT 15827 15870 Plectin 393.
 FT REPEAT 15872 15915 Plectin 394.
 FT REPEAT 15917 15960 Plectin 395.
 FT REPEAT 15962 16005 Plectin 396.
 FT REPEAT 16007 16050 Plectin 397.
 FT REPEAT 16052 16095 Plectin 398.
 FT REPEAT 16097 16140 Plectin 399.
 FT REPEAT 16142 16185 Plectin 400.
 FT REPEAT 16187 16230 Plectin 401.
 FT REPEAT 16232 16275 Plectin 402.
 FT REPEAT 16277 16320 Plectin 403.
 FT REPEAT 16322 16365 Plectin 404.
 FT REPEAT 16367 16410 Plectin 405.
 FT REPEAT 16412 16455 Plectin 406.
 FT REPEAT 16457 16500 Plectin 407.
 FT REPEAT 16502 16545 Plectin 408.
 FT REPEAT 16547 16590 Plectin 409.
 FT REPEAT 16592 16635 Plectin 410.
 FT REPEAT 16637 16680 Plectin 411.
 FT REPEAT 16682 16725 Plectin 412.
 FT REPEAT 16727 16770 Plectin 413.
 FT REPEAT 16772 16815 Plectin 414.
 FT REPEAT 16817 16860 Plectin 415.
 FT REPEAT 16862 16905 Plectin 416.
 FT REPEAT 16907 16950 Plectin 417.
 FT REPEAT 16952 16995 Plectin 418.
 FT REPEAT 16997 17040 Plectin 419.
 FT REPEAT 17042 17085 Plectin 420.
 FT REPEAT 17087 17130 Plectin 421.
 FT REPEAT 17132 17175 Plectin 422.
 FT REPEAT 17177 17220 Plectin 423.
 FT REPEAT 17222 17265 Plectin 424.
 FT REPEAT 17267 17310 Plectin 425.
 FT REPEAT 17312 17355 Plectin 426.
 FT REPEAT 17357 17400 Plectin 427.
 FT REPEAT 17402 17445 Plectin 428.
 FT REPEAT 17447 17490 Plectin 429.
 FT REPEAT 17492 17535 Plectin 430.
 FT REPEAT 17537 17580 Plectin 431.
 FT REPEAT 17582 17625 Plectin 432.
 FT REPEAT 17627 17670 Plectin 433.
 FT REPEAT 17672 17715 Plectin 434.
 FT REPEAT 17717 17760 Plectin 435.
 FT REPEAT 17762 17805 Plectin 436.
 FT REPEAT 17807 17850 Plectin 437.
 FT REPEAT 17852 17895 Plectin 438.
 FT REPEAT 17897 17940 Plectin 439.
 FT REPEAT 17942 17985 Plectin 440.
 FT REPEAT 17987 18030 Plectin 441.
 FT REPEAT 18032 18075 Plectin 442.
 FT REPEAT 18077 18120 Plectin 443.
 FT REPEAT 18122 18165 Plectin 444.
 FT REPEAT 18167 18210 Plectin 445.
 FT REPEAT 18212 18255 Plectin 446.
 FT REPEAT 18257 18300 Plectin 447.
 FT REPEAT 18302 18345 Plectin 448.
 FT REPEAT 18347 18390 Plectin 449.
 FT REPEAT 18392 18435 Plectin 450.
 FT REPEAT 18437 18480 Plectin 451.
 FT REPEAT 18482 18525 Plectin 452.
 FT REPEAT 18527 18570 Plectin 453.
 FT REPEAT 18572 18615 Plectin 454.
 FT REPEAT 18617 18660 Plectin 455.
 FT REPEAT 18662 18705 Plectin 456.
 FT REPEAT 18707 18750 Plectin 457.
 FT REPEAT 18752 18795 Plectin 458.
 FT REPEAT 18797 18840 Plectin 459.
 FT REPEAT 18842 18885 Plectin 460.
 FT REPEAT 18887 18930 Plectin 461.
 FT REPEAT 18932 18975 Plectin 462.
 FT REPEAT 18977 19020 Plectin 463.
 FT REPEAT 19022 19065 Plectin 464.
 FT REPEAT 19067 19110 Plectin 465.
 FT REPEAT 19112 19155 Plectin 466.
 FT REPEAT 19157 19200 Plectin 467.
 FT REPEAT 19202 19245 Plectin 468.
 FT REPEAT 19247 19290 Plectin 469.
 FT REPEAT 19292 19335 Plectin 470.
 FT REPEAT 19337 19380 Plectin 471.
 FT REPEAT 19382 19425 Plectin 47

```

DR EMEL; AB076670; BAC00814.1; -
DR HSP; P32851; 1SFC
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:000865; F:protein transporter activity; IEA.
DR GO; GO:000686; F:intracellular protein transport; IEA.
DR InterPro; IPR006012; Syntaxin.
DR InterPro; IPR006011; Syntaxin_N.
DR InterPro; IPR010989; t-snare.
DR InterPro; IPR000727; T_SNARE.
DR Pfam; PF05739; SNARE; I.
DR Pfam; PF00804; Syntaxin; 1.
DR SMART; SM00503; SynN; 1.
DR SMART; SM00397; t_SNARE; 1.
DR PROSITE; PS00914; SYNTAXIN; 1.
DR PROSITE; PS0192; T_SNARE; 1.
DR SEQUENCE 285 AA, 33026 MW, 5FB1D6B10B77DEAB CRC64;
SQ
Query Match 33.8%; Score 65.5; DB 2; Length 286;
Best Local Similarity 39.5%; Pred. No. 10;
Matches 17; Conservative 6; Mismatches 13; Indels 7; Gaps
QY 5 PYGRDAETLQKQETIK-----AFLLKLEALMASNDNAKNT 40
DB 65 PSGRTELEELNEEIKKIANKIRARLKATIQSFQAQSENAVRT 107
| | | | | : | | | | | | | | | | | | | | |
RESULT 14
GCC2 MOUSE STANDARD; PRT; 1679 AA.
ID GCC2_MOUSE Q8BR44; Q8RQ05; Q9CT45;
AC Q8CHG3; Q8BR44; Q8RQ05; Q9CT45;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE GRIP and coiled-coil domain-containing protein 2 (Golgi coiled coil
DE protein GGC185).
GN Name=Gcc2; Synonyms=Kiaa0336;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE OF 1-769 FROM N.A.
RC TISSUE=Breast cancer;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603999;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bustov K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton W., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshuyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J.J., Schmutz J.J., Myers R.N.,
RA Bateman A., Birney D., Bolintineanu D.S., Botstein D., Bunick D.,
RA Bucherfeld V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE OF 1-442 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
RX MEDLINE=N22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakado I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Isagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D., Burt C., Quackenbush J.,
RA Schriml L.M., Kapin A., Matsuda H., Batalov S., Beise K.W.,

```

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.B., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE OF 49-1679 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara O., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48 (2003).
CC -!- FUNCTION: Probably involved in maintaining Golgi structure (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; peripheral membrane protein
CC associated with the trans-Golgi network (By similarity).
CC -!- DOMAIN: Extended rod-like protein with coiled-coil domains.
CC -!- SIMILARITY: Belongs to the golgin family.
CC -!- SIMILARITY: Contains 1 GRIP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC027339; AAH27339.1; ALT_INIT.
DR EMBL; AK011206; BAB27466.2; -.
DR EMBL; AK045701; BAC32463.1; -.
DR EMBL; AB093232; BAC41416.1; -.
DR PIR; PTC0649; PT0712.
DR HSP; P05354; INWO.
DR MGD; MGI:1917547; 2600014C01Rik.
DR InterPro; IPR000237; GRIP.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF01465; GRIP; 1.
DR PROSITE; PS0913; GRIP; 1.
KW Coiled coil; Golgi stack.
FT DOMAIN 31 1613 Coiled coil (Potential).
FT DOMAIN 83 501 Glu-rich.
FT DOMAIN 1604 1654 GRIP.
FT CONFLICT 440 440 I -> M (in Ref. 2).
SQ SEQUENCE 1679 AA; 194443 MW; 6E2F84263B5E0103 CRC64;

Query Match 33.8%; Score 65.5; DB 1; Length 1679;
Best Local Similarity 39.0%; Pred. No. 59;
Matches 16; Conservative 6; Mismatches 14; Indels 5; Gaps 1;

QY 4 APVGRDAETIQKQ-----KETIKAFKLEALMASNDNANK 39
DB 694 AQLSRDLALQAQDPAHKEHVAEPFKLQLMVERDDLNK 734
RESULT 15
Q82X23 PRELIMINARY; PRT; 636 AA.
ID Q82X23;
AC Q82X23;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Atpase components of ABC transporters with duplicated Atpase
DE domains.
GN OrderedLocusNames=NE0477;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; BX321857; CAD84388.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti...; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 636 AA; 71849 MW; 2C822331AEC0288 CRC64;
Query Match 33.8%; Score 65; DB 2; Length 636;
Best Local Similarity 36.4%; Pred. No. 26;
Matches 12; Conservative 11; Mismatches 10; Indels 0; Gaps 0;
QY 3 MAPVGRDAETIQKQETIKAFKLEALMASND 35
DB 555 LAPRGHTDTLEKQLDNIOKLOALJETILADNN 587
Search completed: November 10, 2004, 14:50:14
Job time : 55.283 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 12.0792 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-53

Perfect score: 162

Sequence: 1 CREQAEITGLRLASLGKFNKIVHSSMTRAET 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	35.2	248	4	US-09-107-532A-4169
2	56	34.6	178	4	US-09-134-000C-4180
3	53	32.7	250	4	US-09-252-991A-30269
4	51	31.5	90	4	US-09-248-796A-24252
5	51	31.5	242	4	US-09-107-532A-5265
6	48.5	29.9	122	4	US-09-328-352-6864
7	48.5	29.9	355	3	US-09-000-040-3
8	48.5	29.9	355	3	US-09-000-084-2
9	47	29.0	316	2	US-08-846-762-9
10	47	29.0	392	4	US-09-252-991A-20001
11	47	29.0	445	4	US-09-252-991A-28655
12	47	29.0	540	4	US-09-914-259-30
13	46	28.4	143	3	US-09-439-313-483
14	46	28.4	143	4	US-09-636-215-483
15	46	28.4	143	4	US-09-685-166A-483
16	46	28.4	143	4	US-09-679-426-483
17	46	28.4	194	4	US-09-134-000C-3470
18	46	28.4	347	4	US-09-248-796A-15247
19	45.5	28.1	236	1	US-08-442-063A-42
20	45.5	28.1	282	1	US-08-442-063A-45
21	45.5	28.1	307	1	US-08-442-063A-48
22	45.5	28.1	333	1	US-08-442-063A-27
23	45.5	28.1	342	1	US-08-272-919-2
24	45.5	28.1	342	1	US-08-619-916-2
25	45.5	28.1	342	5	PCT-US95-08542-2
26	45.5	28.1	353	6	5340934
27	45.5	28.1	359	1	US-08-303-238-4

28	45.5	28.1	359	3	US-08-458-834-4	Sequence 4, Appli
29	45.5	28.1	359	4	US-09-538-092-868	Sequence 868, App
30	45	27.8	323	4	US-09-489-039A-8638	Sequence 8638, Ap
31	45	27.8	559	1	US-08-313-553-15	Sequence 15, Appl
32	45	27.8	559	3	US-08-767-933-15	Sequence 15, Appl
33	45	27.8	800	4	US-09-248-796A-19106	Sequence 19106, A
34	44.5	27.5	281	4	US-09-538-092-652	Sequence 652, App
35	44.5	27.5	545	4	US-09-270-767-46738	Sequence 46738, A
36	44	27.2	191	4	US-09-270-767-32702	Sequence 32702, A
37	44	27.2	191	4	US-09-270-767-47919	Sequence 47919, A
38	44	27.2	224	4	US-09-248-796A-17090	Sequence 17090, A
39	44	27.2	236	4	US-09-134-000C-5155	Sequence 5155, Ap
40	44	27.2	270	4	US-09-071-035-186	Sequence 186, App
41	44	27.2	292	4	US-09-071-035-186	Sequence 186, App
42	44	27.2	415	3	US-09-080-044-6	Sequence 6, Appli
43	44	27.2	415	3	US-09-531-857A-6	Sequence 6, Appli
44	44	27.2	497	3	US-09-134-001C-5114	Sequence 5114, Ap
45	44	27.2	1049	3	US-08-772-270A-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-107-532A-4169
; Sequence 4169, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4169:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...248
; SEQUENCE DESCRIPTION: SEQ ID NO: 4169:
US-09-107-532A-4169


```

Query Match      29.9%; Score 48.5; DB 3; Length 355;
Best Local Similarity 44.8%; Pred.No.12; Mismatches 1; Gaps 1;
Matches 13; Conservative 7; Indels 8; Indels 1; Gaps 1;

QY          5 AELTGLRLASLGL-KFNKIVHSSMTRAIE 32
Db          271 AIVTNVRMSFVELNQFAKVQHSHMRGIQ 299
              |||::|||:|||||::|||:
RESULT 8
US-09-000-064-2
; Sequence 2, Application US/09000064
; Patent No. 6214592
; GENERAL INFORMATION:
; APPLICANT: Crouzet, Joel
; APPLICANT: Favre-Bulle, Olivier
; APPLICANT: Jourdat, Catherine
; APPLICANT: Le Cog, Anne-Marie
; APPLICANT: Petre, Dominique
; TITLE OF INVENTION: ENZYMES AND MICROORGANISMS HAVING AMIDASE ACTIVITY FOR
; FILE REFERENCE: 003025-036
; CURRENT APPLICATION NUMBER: US/09/000,064
; CURRENT FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: PCT/FR96/01119
; EARLIER FILING DATE: 1996-07-17
; EARLIER APPLICATION NUMBER: FR 9508916
; EARLIER FILING DATE: 1995-07-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 355
; TYPE: PRP
; ORGANISM: Escherichia coli
US-09-000-064-2

Query Match      29.9%; Score 48.5; DB 3; Length 355;
Best Local Similarity 44.8%; Pred.No.12; Mismatches 1; Gaps 1;
Matches 13; Conservative 7; Indels 8; Indels 1; Gaps 1;

QY          5 AELTGLRLASLGL-KFNKIVHSSMTRAIE 32
Db          271 AIVTNVRMSFVELNQFAKVQHSHMRGIQ 299
              |||::|||:|||||::|||:
RESULT 9
US-08-846-762-9
; Sequence 9, Application US/08846762A
; Patent No. 5934072
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Burrows, Lori
; APPLICANT: Charter, Deborah
; APPLICANT: de Kievit, Teresa
; TITLE OF INVENTION: No. 5934072el Proteins Involved in the Synthesis and Assembly
; FILE REFERENCE: 6580-089
; CURRENT APPLICATION NUMBER: US/08/846,762A
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 316
; TYPE: PRP
; ORGANISM: Pseudomonas aeruginosa
US-08-846-762-9

Query Match      29.0%; Score 47; DB 2; Length 316;
Best Local Similarity 57.9%; Pred.No.19; Mismatches 5; Indels 0; Gaps 0;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY          10 LRLASLGLKFNKIVHSSMT 28

```

```
Db      116 LAVASLGKINKVLVLSLT 134

RESULT 10
US-09-252-991A-20001
; Sequence 20001, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20001
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20001

Query Match      29.0%; Score 47; DB 4; Length 392;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 14; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy      5 AELTGLASLGKFKNKIVHSSMTTRA 30
Db      359 ALLSGERLASLGKKGQKRVASFDPA 384

RESULT 11
US-09-252-991A-28655
; Sequence 28655, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28655
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28655

Query Match      29.0%; Score 47; DB 4; Length 445;
Best Local Similarity 59.9%; Pred. No. 29;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      10 LRLASLGKFKNKIVHSSMT 28
Db      150 LAVASLGKINKVLVLSLT 168

RESULT 12
US-09-914-259-30
; Sequence 30, Application US/09914259
; Patent No. 6495336
```

```
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-914-259-30

Query Match      29.0%; Score 47; DB 4; Length 540;
Best Local Similarity 38.5%; Pred. No. 36;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy      8 TGLRLASLGKFKNKIVHSSMTTRA1ET 33
Db      11 TGLRTVOQGLEALREHSTISNTLET 36

RESULT 13
US-09-439-313-483
; Sequence 483, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 483
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-439-313-483

Query Match      28.4%; Score 46; DB 3; Length 143;
Best Local Similarity 36.8%; Pred. No. 10;
Matches 14; Conservative 8; Mismatches 8; Indels 8; Gaps 2;

Qy      1 CREQAELTGLRLASL---GLKXF---NKIVHSSMTTRA 30
Db      53 CRQPKLPNSRLSLLPWRDLKFVPRQDKLTRSSVSA 90

RESULT 14
US-09-636-215-483
; Sequence 483, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 38.1679 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-53

Perfect score: 162

Sequence: 1 CREQAELTGLRLASLGLKFNKIVHSSMTRALET 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162	100.0	33	14	US-10-092-750-53
2	153	94.4	255	16	US-10-408-765A-1843
3	58	35.8	121	15	US-10-424-599-263315
4	58	35.8	1059	16	US-10-437-963-147607
5	58	35.8	1112	16	US-10-437-963-147651
6	57	35.2	258	15	US-10-425-114-36853
7	57	35.2	511	14	US-10-310-154-726
8	54	33.3	205	16	US-10-437-963-162719
9	52	32.1	378	15	US-10-424-599-198729
10	51.5	31.8	307	14	US-10-369-493-11212
11	51	31.5	211	14	US-10-369-493-2284
12	51	31.5	444	15	US-10-282-122A-53813
13	51	31.5	1299	16	US-10-437-963-147640

14	50.5	31.2	438	14	US-10-156-761-13404	Sequence 13404, A
15	49	30.2	206	14	US-10-369-493-12082	Sequence 12082, A
16	49	30.2	223	14	US-10-369-493-4610	Sequence 4610, Ap
17	49	30.2	223	14	US-10-369-493-7367	Sequence 7367, Ap
18	49	30.2	268	15	US-10-424-599-164742	Sequence 164742, A
19	49	30.2	308	15	US-10-425-114-60021	Sequence 60021, A
20	49	30.2	311	14	US-10-369-493-316	Sequence 316, App
21	49	30.2	654	15	US-10-425-114-65105	Sequence 65105, A
22	49	30.2	815	17	US-10-425-115-273644	Sequence 273644, A
23	49	30.2	992	17	US-10-739-930-7903	Sequence 7903, Ap
24	49	30.2	1001	17	US-10-425-115-231321	Sequence 231321, A
25	49	30.2	1030	15	US-10-425-114-62748	Sequence 62748, A
26	49	30.2	1080	17	US-10-425-115-321330	Sequence 321330, A
27	48	29.6	57	17	US-10-425-115-329406	Sequence 329406, A
28	48	29.6	237	14	US-10-369-493-16966	Sequence 16966, A
29	48	29.6	261	16	US-10-437-963-176174	Sequence 176174, A
30	48	29.6	505	15	US-10-425-114-64344	Sequence 64344, A
31	48	29.6	834	15	US-10-221-625-34	Sequence 34, Appl
32	48	29.6	848	17	US-10-425-115-276160	Sequence 276160, A
33	48	29.6	850	14	US-10-094-749-2578	Sequence 2578, Ap
34	48	29.6	869	17	US-10-425-115-276161	Sequence 276161, A
35	47.5	29.3	115	17	US-10-425-115-203921	Sequence 203921, A
36	47.5	29.3	306	14	US-10-369-493-1088	Sequence 1088, Ap
37	47.5	29.3	1525	16	US-10-437-963-199176	Sequence 199176, A
38	47	29.0	311	14	US-10-369-493-185	Sequence 185, App
39	47	29.0	316	14	US-10-216-209-9	Sequence 9, Appl
40	47	29.0	320	15	US-10-282-122A-63860	Sequence 63860, A
41	47	29.0	411	9	US-09-815-242-5128	Sequence 5128, Ap
42	47	29.0	411	15	US-10-282-122A-43441	Sequence 43441, A
43	47	29.0	441	15	US-10-112-944-855	Sequence 855, App
44	47	29.0	540	14	US-10-080-608A-30	Sequence 30, Appl
45	47	29.0	540	14	US-10-370-685-119	Sequence 119, App

ALIGNMENTS

RESULT 1
US-10-092-750-53
; Sequence 53, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-53

Query Match 100.0%; Score 162; DB 14; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.1e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CREQAELTGLRLASLGLKFNKIVHSSMTRALET 33

Db 1 CREQAELTGLRLASLGLKFNKIVHSSMTRALET 33

RESULT 2

US-10-408-765A-1843
; Sequence 1843, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1843
 ; LENGTH: 255
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-1843

Query Match 94.4%; Score 153; DB 16; Length 255;
 Best Local Similarity 100.0%; Pred. No. 5.6e-15;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 REQAEITGLRLASLGLKFKNKIVHSSMTAIE 33
 DB 125 REQAEITGLRLASLGLKFKNKIVHSSMTAIE 156

RESULT 3
 US-10-424-599-263315
 ; Sequence 263315, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J
 ; APPLICANT: Kovalic, David K
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 263315
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_79796C.1.pep
 US-10-424-599-263315

Query Match 35.8%; Score 58; DB 15; Length 121;
 Best Local Similarity 43.3%; Pred. No. 0.86;
 Matches 13; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 3 EQAEITGLRLASLGLKFKNKIVHSSMTAIE 32
 DB 2 EDAYANGVRVGSGLGMVNAVILGFVMSLAVE 31

RESULT 4
 US-10-437-963-147651
 ; Sequence 147651, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 147607
 ; LENGTH: 1059
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_4811C.1.pep
 US-10-437-963-147607

Query Match 35.8%; Score 58; DB 16; Length 1059;
 Best Local Similarity 43.8%; Pred. No. 11;
 Matches 14; Conservative 6; Mismatches 10; Indels 2; Gaps 1;

QY 1 CREQA--ELTGLRLASLGLKFKNKIVHSSMTA 30
 DB 652 CRRSGNTFSDKRRKDLVAFNDFVHNSLTRA 683

RESULT 5
 US-10-437-963-147651
 ; Sequence 147651, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 147651
 ; LENGTH: 1112
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1112)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_4815C.1.pep
 US-10-437-963-147651

Query Match 35.8%; Score 58; DB 16; Length 1112;
 Best Local Similarity 43.8%; Pred. No. 12;
 Matches 14; Conservative 6; Mismatches 10; Indels 2; Gaps 1;

QY 1 CREQA--ELTGLRLASLGLKFKNKIVHSSMTA 30
 DB 705 CRRSGNTFSDKRRKDLVAFNDFVHNSLTRA 736

RESULT 6
 US-10-425-114-36853
 ; Sequence 36853, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 36853
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700749341_FLI.pep
US-10-425-114-36853

Query Match 35.2%; Score 57; DB 15; Length 258;
Best Local Similarity 43.3%; Pred. No. 3;
Matches 13; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 3 EQAELTGLRLASLGLKFNKIVHSSMTTRAIE 32
DB 55 EDAYAGVRVGSGLGLMNVAVLGFMSLAVE 94

RESULT 7

US-10-310-154-726
; Sequence 726, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Mollan
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madison, Linda L.
; APPLICANT: Mallory, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padnavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanqun
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang

; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; PRIOR FILING DATE: 2002-12-04
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 726
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
US-10-310-154-726

Query Match 35.2%; Score 57; DB 14; Length 511;
Best Local Similarity 43.3%; Pred. No. 6.7;
Matches 13; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 3 EQAELTGLRLASLGLKFNKIVHSSMTTRAIE 32
DB 318 EDAYAGVRVGSGLGLMNVAVLGFMSLAVE 347

RESULT 8

US-10-437-963-162719
; Sequence 162719, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162719
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61784C.1.pep
US-10-437-963-162719

Query Match 33.3%; Score 54; DB 16; Length 205;
Best Local Similarity 38.7%; Pred. No. 6.6;
Matches 12; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 3 EQAELTGLRLASLGLKFNKIVHSSMTTRAIE 33
DB 33 EEAAGQKLLAAGYEFDAFTSVLTRAINT 63

RESULT 9

US-10-424-599-198729
; Sequence 198729, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198729
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_21476C.1.pap
US-10-424-599-198729

Query Match 32.1%; Score 52; DB 15; Length 378;
Best Local Similarity 40.6%; Pred. No. 27;
Matches 13; Conservative 8; Mismatches 9; Indels 2; Gaps 2;

QY 1 CREQAEITGLRLASGLKFK-NKIVHSSMTRAI 31
Db 171 CNDQADMTG-RLLAGSLTFDNRIMHYIIVRIL 201

RESULT 10
US-10-369-493-11212
; Sequence 11212, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11212
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Methanosarcina mazei
US-10-369-493-11212

Query Match 31.8%; Score 51.5; DB 14; Length 307;
Best Local Similarity 44.0%; Pred. No. 26;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 2 RQAEITGLRLASGLG-LKENKIVHS 25
Db 147 RRESHLEGLKIAMAGDLKRYGRTVHS 171

RESULT 11
US-10-369-493-2284
; Sequence 2284, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2284
; LENGTH: 211
; TYPE: PRT

```

```

; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2284

Query Match 31.5%; Score 51; DB 14; Length 211;
Best Local Similarity 41.9%; Pred. No. 20;
Matches 13; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 3 EQAEITGLRLASGLKFKNKIVHSSMTRAIET 33
Db 40 KEAKJGGERLKSRGYKFDIAFTSALQRAOKT 70

RESULT 12
US-10-282-122A-53813
; Sequence 53813, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 53813
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Corynebacterium diphtheriae
US-10-282-122A-53813

Query Match 31.5%; Score 51; DB 15; Length 444;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 12; Conservative 8; Mismatches 8; Indels 2; Gaps 1;

QY 2 RQAEITGLRLASGLKFKNKIVHSSMTRAI 31
Db 112 RELGEVIGLIATIG--FTLVVESAVAVL 139

RESULT 13
US-10-437-963-147640

```

```
; Sequence 147640, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147640
; LENGTH: 1299
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4814C.1.pap
US-10-437-963-147640

Query Match      31.5%; Score 51; DB 16; Length 1299;
Best Local Similarity 55.0%; Pred. No. 1.7e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy      11 RLASGLKFNKIVHSSMTRA 30
Db      354 RKDLYAEFNDIVHNSLTRA 373

RESULT 14
US-10-156-761-13404
; Sequence 13404, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIEA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, WASHIHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13404
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13404

Query Match      31.2%; Score 50.5; DB 14; Length 438;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 14; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

Qy      2 RQAEITGLRLASGLKFNKIVHSSMTRAET 33
Db      268 RQAEIVGAALAARG-TIQAVVSSPLARCRET 298

RESULT 15
US-10-369-493-12082
; Sequence 12082, Application US/10369493
```

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12082
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12082

Query Match      30.2%; Score 49; DB 14; Length 206;
Best Local Similarity 36.7%; Pred. No. 39;
Matches 11; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy      4 QAEITGLRLASGLKFNKIVHSSMTRAET 33
Db      36 EAKAAGQKLKARGLKFDIAFTSALSRAQKT 65

Search completed: November 11, 2004, 02:43:09
Job time : 39.2179 secs
```

This Page Blank (used),

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 8.09434 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-53
Perfect score: 162
Sequence: 1 CREQAEITGRLASLGKFNKIVHSSMTIAET 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	93	57.4	289	2	T12682	hypothetical prote
2	82	50.6	269	2	T16702	hypothetical prote
3	55	34.0	230	2	S67185	hypothetical prote
4	53	32.7	164	2	AS1834	hypothetical prote
5	53	32.7	183	2	T46943	mannopine synthesi
6	53	32.7	236	2	G83416	hypothetical prote
7	52.5	32.4	331	2	G72273	probable dioxygena
8	52	32.1	227	2	A64091	phosphoglycerate m
9	52	32.1	229	2	AE1350	phosphoglyceromuta
10	52	32.1	229	2	AH1720	phosphoglyceromuta
11	51.5	31.8	98	2	A83508	hypothetical prote
12	51	31.5	211	2	S43214	phosphoglycerate m
13	50	30.9	217	2	AP3540	phosphoglycerate m
14	50	30.9	368	2	A71727	hypothetical prote
15	49.5	30.6	508	2	G84564	probable sugar tra
16	48.5	29.9	355	2	A47039	nylon oligomer-deg
17	48.5	29.9	875	2	F84953	preprotein translo
18	48	29.6	164	1	S74345	phosphohistidine p
19	48	29.6	228	2	C40649	phosphoglycerate m
20	48	29.6	237	2	D87529	phosphoglycerate m
21	48	29.6	252	2	F87259	hypothetical prote
22	48	28.6	1063	2	T38420	probable DNA helic
23	47.5	29.3	306	2	D64497	aspartate carbamoy
24	47	29.0	252	2	B83584	probable molybdenu
25	47	29.0	306	1	OWSEAC	aspartate carbamoy
26	47	29.0	310	2	A84973	aspartate carbamoy
27	47	29.0	320	2	D87085	ornithine carbamoy
28	47	29.0	346	2	T10173	sterol 24-C-methyl
29	47	29.0	411	2	C83251	O-antigen transloc

30 47 29.0 463 2 D97814 chromosomal replic
31 47 29.0 463 2 C71665 chromosomal replic
32 47 29.0 522 2 S41865 kinein light chai
33 47 29.0 537 2 H89052 protein C18C4.10
34 47 29.0 540 2 S41864 kinein light chai
35 47 29.0 563 2 S47997 kinein light chai
36 46.5 28.7 447 2 C96497 glyceraldehyde-3-p
37 46.5 28.7 489 2 F82085 glutamate synthase
38 46.5 28.7 810 2 D64090 glycerol-3-phospha
39 46 28.4 283 2 F90066 pantoate beta-alan
40 46 28.4 303 2 C86011 hypothethical prote
41 46 28.4 303 2 C91165 hypothethical prote
42 46 28.4 311 2 AD1059 aspartate carbamoy
43 46 28.4 370 2 S22124 L-selectin precurs
44 46 28.4 391 2 G90266 na(+)/H(+) antipor
45 46 28.4 411 2 T44155 hypothethical prote

ALIGNMENTS

RESULT 1

T12682
hypothetical protein 63B12.4 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
R;Accession: T12682
R;Ferrat, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, January 1998
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A;Reference number: Z17572
A;Accession: T12682
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-289 <FER>
A;Cross-references: UNIPROT:O46084; EMBL:AL021106; NID:e1371406; PID:e1249575; PIDN:CA7
A;Experimental source: clone cosmid 63B12
C;Genetics:
A;Cross-references: FlyBase:FBgn0000117
A;Introns: 186/3
A;Note: 63B12.4

Query Match 57.4%; Score 93; DB 2; Length 289;
Best Local Similarity 59.4%; Pred. No. 2.4e-06;
Matches 19; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 2 REQAEITGRLASLGKFNKIVHSSMTIAET 33
Db 114 RKQAEFTGKELCELGKWKVASTVRAQET 145

RESULT 2

T16702
hypothetical protein R07G3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
R;Accession: T16702
R;Connell, M.
submitted to the EMBL Data Library, July 1995
A;Description: The sequence of C. elegans cosmid R07G3.
A;Reference number: Z18562
A;Accession: T16702
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-269 <CON>
A;Cross-references: UNIPROT:Q09422; EMBL:U23452; NID:g733564; PID:g733568; PIDN:AAC467
A;Experimental source: strain Bristol N2
C;Genetics:
A;Introns: 62/2; 103/1; 175/3; 220/2

Query Match 50.6%; Score 82; DB 2; Length 269;
Best Local Similarity 56.2%; Pred. No. 0.00011;

R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermoplasma acidophilum*

A; Reference number: A72200; MUID:99287316; PMID:10360571

A; Accession: G72273

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-331 <ARN>

A; Cross-references: UNIPROT:Q9X110; GB:AE001783; GB:AE000512; NID:G4981832; PIDN:AAD3635

A; Experimental source: strain MSB8

C; Genetics:

A; Gene: TM1284

C; Superfamily: probable dioxigenase; Rieske [2Fe-2S] homology

C; Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein

F; 37-86/Domain: Rieske [2Fe-2S] homology <RSK>

F; 47,49,68,71/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status predicted

Query Match 32.4%; Score 52.5; DB 2; Length 331;

Best Local Similarity 43.8%; Pred. No. 4.3;

Matches 14; Conservative 4; Mismatches 9; Indels 5; Gaps 1;

QY 7 LTGLR-----LASGLKFNKIVHSSMTTRAET 33

DB 259 LTGKFLDKLAVLGKPNKIVLHQDKRWVET 290

RESULT 8

AH1720

phosphoglycerate mutase (EC 5.4.2.1) - Haemophilus influenzae (strain Rd KW20)

C; Species: Haemophilus influenzae

C; Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 16-Aug-2004

C; Accession: A64091

R; Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995

A; Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A; Reference number: A64000; MUID:95350630; PMID:7542800

A; Accession: A64091

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-227 <TIGR>

A; Cross-references: UNIPROT:P44865; GB:U32760; GB:L42023; NID:G1573764; PIDN:AAC22416.1

C; Superfamily: Cofactor-dependent phosphoglycerate mutase; phosphoglycerate mutase homol

C; Keywords: intramolecular transferase; isomerase; phosphohistidine; phosphoprotein

F; 3-216/Domain: phosphoglycerate mutase homology <PGMH>

F; 7,59,181/Active site: Arg, Arg, His #status predicted

F; 8/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 32.1%; Score 52; DB 2; Length 227;

Best Local Similarity 38.7%; Pred. No. 3.4;

Matches 12; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 3 EQAELTGLRLASGLKFNKIVHSSMTTRAET 33

DB 33 EEAATAGKRLDKGYEFDVAFVTSVLTTRAIT 63

RESULT 9

AH1350

phosphoglyceromutase 1 homolog lmo2205 [imported] - Listeria monocytogenes (strain EGD-e)

C; Species: Listeria monocytogenes

C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004

C; Accession: AE1350

R; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001

A; Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,

A; Title: Comparative genomics of Listeria species.

A; Reference number: AB1077; MUID:21537279; PMID:11679669

A; Accession: AE1350

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-229 <GLA>

A; Cross-references: UNIPROT:Q8Y571; GB:NC_003210; PIDN:CAD00283.1; PID:G16411675; GSPDB

A; Experimental source: strain EGD-e

C; Genetics:

A; Gene: lmo2205

C; Superfamily: Cofactor-dependent phosphoglycerate mutase; phosphoglycerate mutase hom

Query Match 32.1%; Score 52; DB 2; Length 229;

Best Local Similarity 40.0%; Pred. No. 3.4;

Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 4 QAELTGLRLASGLKFNKIVHSSMTTRAET 33

DB 34 EAATAGKRIKEAGLEFDVAFVTSVLTTRAIT 53

RESULT 10

AH1720

phosphoglyceromutase 1 homolog lin2308 [imported] - Listeria innocua (strain Clip11262)

C; Species: Listeria innocua

C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004

C; Accession: AH1720

R; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001

A; Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of Listeria species.

A; Reference number: AB1077; MUID:21537279; PMID:11679669

A; Accession: AH1720

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-229 <GLA>

A; Cross-references: UNIPROT:Q82908; GB:AL592022; PIDN:CAC97536.1; PID:G16414820; GSPDB:

A; Experimental source: strain Clip11262

C; Genetics:

A; Gene: lin2308

C; Superfamily: Cofactor-dependent phosphoglycerate mutase; phosphoglycerate mutase hom

Query Match 32.1%; Score 52; DB 2; Length 229;

Best Local Similarity 40.0%; Pred. No. 3.4;

Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 4 QAELTGLRLASGLKFNKIVHSSMTTRAET 33

DB 34 EAATAGKRIKEAGLEFDVAFVTSVLTTRAIT 63

RESULT 11

AH3508

hypothetical protein PA1096 [imported] - Pseudomonas aeruginosa (strain PAO1)

C; Species: Pseudomonas aeruginosa

C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C; Accession: AB3508

R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Liadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; LiLory, S.; Olson, M.V. Nature 406, 959-964, 2000

A; Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path

A; Reference number: AB2950; MUID:20437337; PMID:10984043

A; Accession: AB3508

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-98 <STO>

A; Cross-references: UNIPROT:Q914N5; GB:AE004540; GB:AE004091; NID:G9947008; PIDN:AAG04.

A; Experimental source: strain PAO1

C; Genetics:

A:Gene: PA1096

Query Match 31.8%; Score 51.5; DB 2; Length 98;
Best Local Similarity 56.5%; Pred. No. 1.6;
Matches 13; Conservative 2; Mismatches 3; Indels

Qy	1 CREQ-----AELTGLRLASLGLK 18
	: :
Db	70 CROOREAIGAEALTGLORASKGAK 92

RESULT 12

S43214
 S43214
 phosphoglycerate mutase (EC 5.4.2.1) - fission yeast (*Schizosaccharomyces pombe*)
 C/Species: Schizosaccharomyces pombe
 C/Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
 C/Accession: S43369; T38414; S43214
 R/Nairn, J.; Price, N.C.; Fothergill-Gilmore, L.A.; Walker, G.E.; Fothergill, J.E.; Dunb
 Biochem. J. 297, 603-608, 1994
 A/Title: The amino acid sequence of the small monomeric phosphoglycerate mutase from the
 A/Reference number: S43369; MUID:94153336; PMID:8110200
 A/Accession: S43369
 A/Molecule type: mRNA
 A/Residues: 1-211 <NAI>
 A/Cross-references: UNIPROT:P36623; EMBL:X75385; NID:G407526; PIDN:CAA53154.1; PID:G4075
 R/Brown, D.; Churcher, C.M.; Barrel, B.G.; Rajadream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, April 1996
 A/Reference number: Z21792
 A/Accession: T38414
 A/Status: preliminary; translated from GB/EMBL/DDBJ

```

Query Match          31.5%; Score 51; DB 2; Length 211;
Best Local Similarity 41.9%; Pred. No. 4.4;
Matches 13: Conservative 7; Mismatches 11; Indels 0; Gaps 0;

```

```

QY      3  EQAELTGLRLASLGLKFNKIVHSSMTRAJET 33
      :::| | | | | | | | | | | | | | | |
Db      40 KEAKLGGGERLKSRGYKFDIAFTSALORAOKT 70

```

RESULT 13

AF3540
phosphoglycerate mutase (EC 5.4.2.1) [imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 16-Aug-2004
C:Accession: AF3540
C:R1:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
M.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier,
A.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD352; PMID:11756688
A:Accession: AF3540
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <KUR>
A:Cross-references: GB:AE008918; PIN:AAL53489.1; PID:g17984392; GSPDB:GN00191
A:Experimental source: strain 16M

C:Keywords: intramolecular transferase; isomerase

Query Match 30.9%; Score 50; DB 2; Length 217;
Best Local Similarity 40.7%; Pred. No. 6.4;
Matches 11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 4 Q A E L T G L R L A S I G L K F N K I V H S S M T R A 30
 : : | | | : | | | : | : | :
D6 47 E K A A G O R L K A A G L K F D I A Y T S A L S R A 73

RESULT 14

A71727
 Hypothetical protein RP165 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
 C:Accession: A71727
 R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, N.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: A71727
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-368 <AND>
 A:Cross-references: UNIPROT:Q9ZD28; GB:AJ235270; GB:AJ235269; NID:G3860572; PIDN:CAA146
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: RP165
 C:Superfamily: Rickettsia prowazekii hypothetical protein RP165

```
Query Match      30.9%; Score 50; DB 2; Length 368;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

QY 5 AELTGLRLASLGLKFNKIV 23
|:|:| | | | | |
Db 152 ADLTGIELHSRNLNFKNSV 170

RESULT 15

G84564
Probable sugar transporter [imported] - Arabidopsis thaliana
C-Species: Arabidopsis thaliana (mouse-ear cress)
C-Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C-Accession: G84564
R.Rubin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.
M.; Koo, H.; Moffat, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
D.; Nierman, W.C.; Weiss, D.; Nusbaumer, A.J.; Young, R.A.; Shumway-Cook, A.;
Nelson, D.L.; Schaefer, S.M.; Stitt, M.J.; Weitz, J.; Zerkow, R.; Ziegler, P.;
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A'Reference number: A84420; MUTID:20083457; PMID:10617197

Query Match 30.6%; Score 49.5; DB 2; Length 508;
Best Local Similarity 37.5%; Pred. No. 20;
Matches 15; Conservative 9; Mismatches 5; Indels 11; Gaps 2;

QY 5 AELTGLRL----ASLGLKFNKIVHS-----SMTPRAIET 33
: : ||| ||| : : | : : : ||| : |||
Db 407 SEIFPLRLRGAOSIGVAVNRIMNATVMSSELSMTKAIT 446

Search completed: November 10, 2004, 14:52:21
Job time : 8.09434 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 43.9585 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-53

Perfect score: 162

Sequence: 1 CREQAEITGLRLASLGLKFNKIVHSSMTRAET 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162	100.0	34	Q96JB0	Q96JB0 homo sapien
2	153	94.4	255	Q96HS1	Q96HS1 homo sapien
3	152	93.8	282	Q80VY8	Q80VY8 mus musculus
4	152	93.8	287	Q8BM78	Q8BM78 m mus muscu
5	152	93.8	288	Q8BX10	Q8BX10 mus musculus
6	151	93.2	252	Q8CZU2	Q8CZU2 mus musculus
7	112	69.1	268	Q8AVL5	Q8AVL5 xenopus lae
8	112	69.1	278	Q6GL33	Q6GL33 xenopus tro
9	93	57.4	289	Q46084	Q46084 drosophila
10	93	57.4	294	Q7K2X2	Q7K2X2 drosophila
11	93	57.4	294	AAL28156	AAL28156 drosophil
12	82	50.6	284	1 YR05_CAEEL	Q09422 caenorhabdi
13	80	49.4	285	Q7PRB9	Q7PRB9 anopheles g
14	77	47.5	280	Q9W173	Q9W173 drosophila
15	59	36.4	259	Q6GNS4	Q6GNS4 xenopus lae
16	58	35.8	152	Q8Z201	Q8Z201 pyrobaculum
17	58	35.8	156	Q8GSP9	Q8GSP9 glycine max
18	58	35.8	1061	Q7XSA43	Q7XSA43 cryza sativ
19	58	35.8	1114	Q8ET16	Q8ET16 cryza sativ
20	56	34.6	175	Q822R6	Q822R6 enterococcu
21	56	34.6	230	1 GP2A_LACPL	Q8T310 lactobacill
22	55	34.0	213	Q7NLX7	Q7NLX7 gloebacter
23	55	34.0	230	Q12040	Q12040 saccharomyc
24	55	34.0	295	Q811V2	Q811V2 plasmodium
25	54	33.3	227	1 GPNA_HAEDU	Q7V128 haemophilus
26	54	33.3	259	Q6GPF4	Q6GPF4 xenopus lae
27	53.5	33.0	355	Q93MC1	Q93MC1 clostridium
28	53	32.7	164	Q82077	Q82077 anabaena sp
29	53	32.7	178	Q9WWM8	Q9WWM8 anabaena sp
30	53	32.7	183	Q9PWA7	Q9PWA7 agrobacteri
31	53	32.7	208	Q9K466	Q9K466 streptomyce

RESULT 1

Q96JB0 PRELIMINARY; PRT; 34 AA.
AC Q96JB0; (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Bcl-XL-binding protein v68 (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=21293069; PubMed=11283018;
RA Hammond P.W., Alpin J., Rise C.E., Wright M., Kreider B.L.;
RT "In Vitro Selection and Characterization of Bcl-XL-binding Proteins
from a Mix of Tissue-specific mRNA Display Libraries.";
RL J. Biol. Chem. 276:20598-20906(2001).
DR EMBL; AF357523; AAK60627.1; -.
FT NON TER 1
FT NON TER 34
SQ SEQUENCE 34 AA; 3775 MW; 9DCDF609B1D5771 CRC64;

Query Match 100.0%; Score 162; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 9e-17; 0; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 0;

Qy 1 CREQAEITGLRLASLGLKFNKIVHSSMTRAET 33

Db 1 CREQAEITGLRLASLGLKFNKIVHSSMTRAET 33

RESULT 2

Q96HS1 PRELIMINARY; PRT; 255 AA.
AC Q96HS1; (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein MGC5352.
GN Name=MGC5352;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008196; AH08196.1; -
DR GO: GO:0003824; P: catalytic activity; IEA.
DR GO: GO:0008152; P: metabolism; IEA.
DR InterPro: IPR001345; PG/BPGM_mutase.
DR Pfam: PF00300; PGAM; 1.
DR Hypothetical protein.
SQ SEQUENCE 255 AA; 28006 MW; 07B881C338BE033C CRC64;

Query Match 94.4%; Score 153; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 REQAEITGLRLASLGLKFNKIVHSSMTIAET 33
DB 125 REQAEITGLRLASLGLKFNKIVHSSMTIAET 156

RESULT 3
Q80VY8 PRELIMINARY; PRT; 282 AA.
AC Q80VY8;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2004 (T-EMBLrel. 24, Last sequence update)
DE 2610528A17Rik protein (fragment).
DE Name=2610528A17Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008196; AH08196.1; -
DR GO: GO:0003824; P: catalytic activity; IEA.
DR GO: GO:0008152; P: metabolism; IEA.
DR InterPro: IPR001345; PG/BPGM_mutase.
DR Pfam: PF00300; PGAM; 1.
DR Hypothetical protein.
SQ SEQUENCE 255 AA; 28006 MW; 07B881C338BE033C CRC64;

Query Match 94.4%; Score 153; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 REQAEITGLRLASLGLKFNKIVHSSMTIAET 33
DB 125 REQAEITGLRLASLGLKFNKIVHSSMTIAET 156

RESULT 4
Q8BM78 PRELIMINARY; PRT; 287 AA.
AC Q8BM78;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Mus musculus 12 days embryo embryonic body between diaphragm region
DE and neck cDNA, RIKEN full-length enriched library, clone:9430011H13
DE product:hypothetical phosphoglycerate mutase family containing
DE protein, full insert sequence.
DE Name=2610528A17Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RA Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RA Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Embryonic body between diaphragm region and neck;
RX MEDLINE=20459374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

```



```

RESULT 6
Q9CZU2
ID Q9CZU2 PRELIMINARY; PRT; 252 AA.
AC Q9CZU2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610528A17 product:hypothetical
DE Phosphoglycerate mutase family containing protein, full insert
DE sequence.
OS Name=2610528A17rik;
GN Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=2030913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Kitsumai T., Tashiro H., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Ikegami T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsushita S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito K., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

```

```

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK012159; BAB28067.1; -
DR MGD; MGI:1919792; F:catalytic activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAW; 1.
KW Hypothetical protein.
SQ SEQUENCE 252 AA; 28695 MW; CC659D74D175ACD7 CRC64;

Query Match 93.2%; Score 151; DB 2; Length 252;
Best Local Similarity 93.8%; Pred. No. 3.2e-14;
Matches 30; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 REQALTLGLRLASGLKFKNKIVHSSMTIAET 33
DB 88 REQALTLGLRLASGLKFKNKIVHSSMTIAET 119

RESULT 7
Q8AVL5
ID Q8AVL5 PRELIMINARY; PRT; 268 AA.
AC Q8AVL5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LOC398484 protein (Fragment).
GN Name=LOC398484;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

```

```
DR EMBL; BC041756; AAH41756.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
FT NON_TER 1
SQ SEQUENCE 268 AA; 30564 MW; 980D4568A19742C CRC64;
Query Match
Best Local Similarity 75.1%; Score 112; DB 2; Length 268;
Matches 24; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 REQALTLGLRLASLGLKFNKIVHSSMTTRAIET 33
DB 118 REQALTLGLRLASLGLKFNKIVHSSMTTRAKET 149

RESULT 8
O6GL33 PRELIMINARY; PRT; 278 AA.
AC O6GL33
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettunen M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RL Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074682; AAH74682.1; -.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
KW Hypothetical protein.
SQ SEQUENCE 278 AA; 31210 MW; 67923E297F554639 CRC64;
Query Match
Best Local Similarity 71.9%; Pred. No. 2.7e-08;
Matches 23; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 2 REQALTLGLRLASLGLKFNKIVHSSMTTRAIET 33
DB 115 REQALTLGLRLASLGLKFNKIVHSSMTTRAKET 146
```

```
RESULT 9
O46084 PRELIMINARY; PRT; 289 AA.
AC O46084
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CG14816-PA (EG:63B12.4 protein).
GN Name:EG:63B12.4; ORFNames:CG14816;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Rosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang Z.Y.,
RA Williams S.M., Woodagat, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu H.O.,
RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
```

RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RL a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

RA [4]
 RP SEQUENCE FROM N.A.

RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Whitfield E.J., Bayraktaroglu L., Berman S.P.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., de Grey A.D., Drysdale R.A.,
 RA Battencourt B.R., Celniker S.E., Russo S., Schroeder A.J., Shu S.Q.,
 RA Harris N.L., Richter J., Russo S., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

RA [5]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RA [6]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

RA [7]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [8]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [9]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [10]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [11]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [12]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [13]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [14]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [15]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [16]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [17]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [18]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA [19]

RP SEQUENCE FROM N.A.

RX FLYBASE;
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RX EMBL; AY060608; AAL28156.2; -;
 DR InterPro; IPR001345; PG/BPGM_mutase.

DR Pfam; PF00300; PGAM; 1.

DR NON_TER

FT SEQUENCE

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

Query Match

57.4%; Score 93; DB 2; Length 294;

Best Local Similarity

59.4%; Pred. No. 2.1e-05;

Matches 19; Conservative

6; Mismatches 7; Indels 0; Gaps 0;

294 AA; 33625 MW; 57323DC668120198 CRC64;

```

RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23452; AAC1545.2; -
DR F01; T16702; T16702.
DR WormPep; R07G3.5; CE31979.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 32490 MW; 820D56CE79E71F90 CRC64;

Query Match 50.6%; Score 82; DB 1; Length 284;
Best Local Similarity 56.2%; Pred. No. 0.00094;
Matches 18; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 REQALTGRLASLGKFNKIVHSSMTTAEIT 33
Db ||||| :||| :||| :||| :|||
119 REQALLGRLANSIDKFTNMTMTVRATET 150

RESULT 13
Q7PRB9 PRELIMINARY; PRT; 285 AA.
AC Q7PRB9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000022774.
GN Name=ENSANGG00000021127;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA0100859; EAA07503.2; -
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
SQ SEQUENCE 285 AA; 32423 MW; C9B01EF80B5859F7 CRC64;

Query Match 49.4%; Score 80; DB 2; Length 285;
Best Local Similarity 56.2%; Pred. No. 0.0019;
Matches 18; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 REQALTGRLASLGKFNKIVHSSMTTAEIT 33
Db ||||| :||| :||| :||| :|||
123 RQQAALSGRLKQALPYDEIVRSTMTTAEIT 154

RESULT 14
Q9WI73 PRELIMINARY; PRT; 280 AA.
AC Q9WI73;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG15874-PA.
GN OREFNames=CG15874;
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
FX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brotter P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mccarthy C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun: release 3 of the Drosophila
RA melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.,
RA "The transposable elements of the Drosophila melanogaster euchromatin:
RA a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield J.E., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Asburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB003464; AA247204.1; -;
DR Flybase; F590035004; CG15874.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
SQ SEQUENCE 280 AA; 31395 MW; 5AAEF130DF1AAC59 CRC64;

Query Match 47.5%; Score 77; DB 2; Length 280;
Best Local Similarity 53.1%; Pred. No. 0.0053;
Matches 17; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 REQAEITGLRLASGLKFKNKIVHSSMTTRAET 33
DB 103 RRQAERTGQRLREMGSLSDHVVASTPRAEET 134

RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield J.E., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Asburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB003464; AA247204.1; -;
DR Flybase; F590035004; CG15874.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001345; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
SQ SEQUENCE 280 AA; 31395 MW; 5AAEF130DF1AAC59 CRC64;

Query Match 47.5%; Score 77; DB 2; Length 280;
Best Local Similarity 53.1%; Pred. No. 0.0053;
Matches 17; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 REQAEITGLRLASGLKFKNKIVHSSMTTRAET 33
DB 103 RRQAERTGQRLREMGSLSDHVVASTPRAEET 134

RESULT 15
Q6GNS4 PRELIMINARY; PRT; 259 AA.
AC Q6GNS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maiz M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC073429; AAH73429.1; -;
DR InterPro; IPR005952; PG/BPGM_mutase.
DR Pfam; PF00300; PGAM; 1.
DR TIGRFAMs; TIGR01258; pgm_1; 1.
DR PROSITE; PS00175; PG_MUTASE; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 29622 MW; FEC00BD6627EAF96 CRC64;

Query Match 36.4%; Score 59; DB 2; Length 259;
Best Local Similarity 38.7%; Pred. No. 2.5;
Matches 12; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 3 EQAELTGLRLASGLKFKNKIVHSSMTTRAET 33
DB 36 KEAECEGKGLKSLGFFDLVFTSLRSIQT 66

Search completed: November 10, 2004, 14:50:16
Job time : 45.9585 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-54

Perfect score: 123
Sequence: 1 GTRISDMKLIADTWQRCNCPA 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.psp:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.psp:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.psp:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.psp:*
5: /cgn2_6/ptodata/1/iaa/FACTUS.COMB.psp:*
6: /cgn2_6/ptodata/1/iaa/backfiles.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	41.5	325	4 US-09-461-325-141	Sequence 141, App
2	51	41.5	325	4 US-10-012-542-141	Sequence 141, App
3	51	41.5	325	4 US-10-115-123-141	Sequence 141, App
4	50	40.7	498	4 US-09-489-039A-13251	Sequence 13251, A
5	47	38.2	1151	3 US-08-836-134-23	Sequence 23, Appl
6	47	38.2	1151	4 US-09-493-784-23	Sequence 23, Appl
7	47	38.2	1232	3 US-08-836-134-2	Sequence 2, Appl
8	47	38.2	1232	4 US-09-493-784-2	Sequence 2, Appl
9	47	38.2	1295	4 US-09-703-872-3	Sequence 3, Appl
10	47	38.2	1403	4 US-09-705-872-1	Sequence 1, Appl
11	45	36.6	269	4 US-09-270-767-39447	Sequence 39447, A
12	45	36.6	269	4 US-09-270-767-54664	Sequence 54664, A
13	44	35.8	153	4 US-09-252-991A-21557	Sequence 21557, A
14	44	35.8	230	4 US-09-248-796A-21362	Sequence 21362, A
15	44	35.8	2476	2 US-08-276-967-2	Sequence 2, Appl
16	43.5	35.4	63	4 US-09-497-491-43	Sequence 43, Appl
17	43.5	35.4	64	4 US-09-497-491-41	Sequence 41, Appl
18	42	34.1	110	4 US-09-847-524-6	Sequence 6, Appl
19	42	34.1	118	4 US-09-270-767-43287	Sequence 43287, A
20	42	34.1	159	4 US-09-252-991A-20138	Sequence 20138, A
21	42	34.1	171	4 US-09-540-236-3559	Sequence 3559, AP
22	42	34.1	177	4 US-09-252-991A-22998	Sequence 22998, A
23	42	34.1	321	3 US-08-915-795-9	Sequence 9, Appl
24	42	34.1	321	4 US-09-847-524-4	Sequence 4, Appl
25	42	34.1	321	4 US-09-296-275-9	Sequence 9, Appl
26	42	34.1	345	4 US-09-270-767-46203	Sequence 46203, A
27	42	34.1	358	3 US-08-915-795-8	Sequence 8, Appl

28 42 34.1 358 4 US-09-847-524-2 Sequence 2, Appli
29 42 34.1 358 4 US-09-296-275-8 Sequence 8, Appli
30 42 34.1 358 4 US-09-438-046-15 Sequence 15, Appl
31 42 34.1 597 4 US-09-252-991A-23152 Sequence 23152, A
32 42 34.1 1213 4 US-09-543-681A-6478 Sequence 6478, AP
33 41 33.3 36 4 US-09-270-767-38053 Sequence 38053, A
34 41 33.3 36 4 US-09-270-767-53270 Sequence 53270, A
35 41 33.3 109 3 US-09-469-186-1 Sequence 1, Appli
36 41 33.3 109 4 US-09-469-186-1 Sequence 1, Appli
37 41 33.3 109 4 US-10-100-037-1 Sequence 1, Appli
38 41 33.3 158 4 US-09-252-991A-26807 Sequence 26807, A
39 41 33.3 197 4 US-09-431-888-8 Sequence 8, Appli
40 41 33.3 212 4 US-08-328-352-7485 Sequence 7485, AP
41 41 33.3 243 4 US-09-252-991A-19853 Sequence 19853, A
42 41 33.3 267 4 US-09-489-039A-7863 Sequence 7863, AP
43 41 33.3 325 3 US-08-915-795-3 Sequence 3, Appli
44 41 33.3 325 4 US-09-296-275-3 Sequence 3, Appli
45 41 33.3 354 3 US-08-915-795-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-461-325-141
; Sequence 141, Application US/09461325A
; Patent No. 8475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 141
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-141

Query Match 41.5%; Score 51; DB 4; Length 325;
Best Local Similarity 47.6%; Pred. No. 3.3;
Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy 1 GTRISDMKLIADTWQRCNCP 21
Db 2 GSVSSMKLAL-----QNC 18

RESULT 2
US-10-012-542-141
; Sequence 141, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/10/012,542


```

; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 141
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-141

```

```

Query Match 41.5%; Score 51; DB 4; Length 325;
Best Local Similarity 47.6%; Pred. No. 3.3;
Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

```

```

Qy 1 GTRISDMLKLIADTWQNCPP 21
Db 2 GSQVSSMLKLAL----QNCPP 18

```

```

RESULT 3
US-10-115-123-141
; Sequence 141, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 141
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-123-141

```

```

Query Match 41.5%; Score 51; DB 4; Length 325;
Best Local Similarity 47.6%; Pred. No. 3.3;
Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

```

```

Qy 1 GTRISDMLKLIADTWQNCPP 21
Db 2 GSQVSSMLKLAL----QNCPP 18

```

```

RESULT 4
US-09-489-039A-13251
; Sequence 13251, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13251
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13251

```

```

Query Match 40.7%; Score 50; DB 4; Length 498;
Best Local Similarity 52.6%; Pred. No. 7.6;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

```

```

Qy 2 TRIS--DMLKLIADTWQNR 18
Db 78 TRLSFEQLKLVSDWQNR 96

```

```

RESULT 5
US-08-836-134-23
; Sequence 23, Application US/08836134A
; Patent No. 6020127
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie
; APPLICANT: Ikeda, Joh-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6020127
; TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/08/836,134A
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-836-134-23

```

```

Query Match 38.2%; Score 47; DB 3; Length 1151;
Best Local Similarity 53.3%; Pred. No. 56;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 7 MLKLIADTWQNCPP 21
Db 488 LLKKIATFLWASGCCP 502

```

```

RESULT 6
US-05-493-784-23
; Sequence 23, Application US/09493784
; Patent No. 6429011
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie

```



```
; APPLICANT: Ikeda, Joh-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6429011
; TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/09/493,784
; PRIOR FILING DATE: 2000-01-28
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-493-784-23

  Query Match      38.2%; Score 47; DB 4; Length 1151;
  Best Local Similarity 53.3%; Pred. No. 56;
  Matches      8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      7 MLKLIADTWORNCPP 21
Db      488 LKKIAFLWASGCCP 502

RESULT 7
US-08-836-134-2
; Sequence 2, Application US/08836134A
; Patent No. 6020127
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie
; APPLICANT: Ikeda, Joh-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6020127
; TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/08/836,134A
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-836-134-2

  Query Match      38.2%; Score 47; DB 3; Length 1232;
  Best Local Similarity 53.3%; Pred. No. 60;
  Matches      8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      7 MLKLIADTWORNCPP 21
Db      479 LKKIAFLWASGCCP 493

RESULT 8
US-09-493-784-2
; Sequence 2, Application US/09493784
; Patent No. 6429011
; GENERAL INFORMATION:
; APPLICANT: Mackenzie, Alex E.
; APPLICANT: Korneluk, Robert G.
; APPLICANT: Mahadevan, Mani S.
; APPLICANT: McLean, Michael
; APPLICANT: Roy, Natalie
; APPLICANT: Ikeda, Joh-e
; TITLE OF INVENTION: Neuronal Apoptosis Inhibitor Protein, Gene Sequence and
; Patent No. 6429011
```

```
; TITLE OF INVENTION: Mutations Causative of Spinal Muscular Atrophy
; FILE REFERENCE: 3477-112, 033477/139914
; CURRENT APPLICATION NUMBER: US/09/493,784
; PRIOR FILING DATE: 2000-01-28
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-493-784-2

  Query Match      38.3%; Score 47; DB 4; Length 1232;
  Best Local Similarity 53.3%; Pred. No. 60;
  Matches      8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      7 MLKLIADTWORNCPP 21
Db      479 LKKIAFLWASGCCP 493

RESULT 9
US-09-705-872-3
; Sequence 3, Application US/09705872
; Patent No. 6617429
; GENERAL INFORMATION:
; APPLICANT: Joh-E IKEDA
; APPLICANT: Kenji YAMAMOTO
; TITLE OF INVENTION: APOPTOSIS INHIBITORY PROTEIN, GENE ENCODING THE PROTEIN
; TITLE OF INVENTION: AND CDNA THEREOF
; FILE REFERENCE: 2000-1110/LC/00653
; CURRENT APPLICATION NUMBER: US/09/705,872
; CURRENT FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: 09/239,797
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-705-872-3

  Query Match      38.2%; Score 47; DB 4; Length 1295;
  Best Local Similarity 53.3%; Pred. No. 63;
  Matches      8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      7 MLKLIADTWORNCPP 21
Db      479 LKKIAFLWASGCCP 493

RESULT 10
US-09-705-872-1
; Sequence 1, Application US/09705872
; Patent No. 6617429
; GENERAL INFORMATION:
; APPLICANT: Joh-E IKEDA
; APPLICANT: Kenji YAMAMOTO
; TITLE OF INVENTION: APOPTOSIS INHIBITORY PROTEIN, GENE ENCODING THE PROTEIN
; TITLE OF INVENTION: AND CDNA THEREOF
; FILE REFERENCE: 2000-1110/LC/00653
; CURRENT APPLICATION NUMBER: US/09/705,872
; CURRENT FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: 09/239,797
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1403
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-705-872-1

Query Match      38.2%; Score 47; DB 4; Length 1403;
Best Local Similarity 53.3%; Pred. No. 69;
Matches      8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      7 MLKLIADTWQRNCCP 21
      :|||:|||:|||
Db      479 LLKIAFLWASGCCP 493

RESULT 11
US-09-270-767-39447
; Sequence 39447, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39447
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39447

Query Match      36.6%; Score 45; DB 4; Length 269;
Best Local Similarity 50.0%; Pred. No. 24;
Matches      9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 GTRISDMKLXIADTWQRN 18
      |||:|||:|||:|||
Db      207 GTNISXIVLDIADIWKRN 224

RESULT 12
US-09-270-767-54664
; Sequence 54664, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54664
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54664

Query Match      36.6%; Score 45; DB 4; Length 269;
Best Local Similarity 50.0%; Pred. No. 24;
Matches      9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 GTRISDMKLXIADTWQRN 18
      |||:|||:|||:|||
Db      207 GTNISXIVLDIADIWKRN 224

RESULT 13
US-09-252-991A-21557
; Sequence 21557, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21557
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21557

Query Match      35.8%; Score 44; DB 4; Length 153;
Best Local Similarity 60.0%; Pred. No. 19;
Matches      6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      12 ADTWQRNCCP 21
      |||:|||:|||
Db      40 AACWRRSCCP 49

RESULT 14
US-09-248-796A-21362
; Sequence 21362, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21362
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-21362

Query Match      35.8%; Score 44; DB 4; Length 230;
Best Local Similarity 50.0%; Pred. No. 30;
Matches      9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 GTRISDMKLXIADTWQRN 18
      |||:|||:|||:|||
Db      105 GTRESDLKLNTDELN 122

RESULT 15
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5651817
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garbers, David L.
; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
; TITLE OF INVENTION: Sperm
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
```

STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,967
FILING DATE: Submitted Herewith
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSD:418\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-967-2

Query Match 35.8%; Score 44; DB 2; Length 2476;
Best Local Similarity 54.5%; Pred. No. 3.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 KLIADTWORNC 19
Db 1376 KEVADAWKNC 1386

Search completed: November 10, 2004, 14:55:39
Job time : 8.10283 secs

This Page Blank (user),

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 25.4453 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-54

Perfect score: 123
Sequence: 1 GTRISDMKLADIATWQRNCCPA 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US12_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	123	100.0	22	14 US-10-092-750-54	Sequence 54, Appl
2	51	41.5	97	9 US-09-864-761-39339	Sequence 39339, A
3	51	41.5	325	14 US-10-012-542-141	Sequence 141, App
4	51	41.5	325	14 US-10-115-123-141	Sequence 141, App
5	51	41.5	525	14 US-10-257-022-11	Sequence 11, Appl
6	51	41.5	581	9 US-09-835-270-2	Sequence 2, Appli
7	51	41.5	581	16 US-10-408-765A-904	Sequence 904, App
8	51	41.5	581	16 US-10-755-889-204	Sequence 204, App
9	50	40.7	490	14 US-10-369-493-766	Sequence 766, App
10	49.5	40.2	332	15 US-10-425-114-68020	Sequence 68020, A
11	49.5	40.2	338	15 US-10-424-599-218174	Sequence 218174,
12	48	39.0	677	9 US-09-925-301-1295	Sequence 1295, Ap
13	48	39.0	1698	14 US-10-080-943-4	Sequence 4, Appli

14 47 38.2 61 15 US-10-424-599-215390 Sequence 215390,
15 47 38.2 203 14 US-10-029-386-33933 Sequence 33933, A
16 47 38.2 486 14 US-10-055-475-12 Sequence 12, Appl
17 47 38.2 487 15 US-10-108-260A-4515 Sequence 4515, Ap
18 47 38.2 729 17 US-10-478-516-18 Sequence 18, Appl
19 47 38.2 782 9 US-09-841-739-9 Sequence 9, Appli
20 47 38.2 782 14 US-10-449-315-9 Sequence 9, Appli
21 47 38.2 1403 8 US-08-913-322-22 Sequence 22, Appl
22 47 38.2 1403 8 US-08-913-322-24 Sequence 24, Appl
23 47 38.2 1403 14 US-10-285-408-1 Sequence 1, Appli
24 47 38.2 1689 14 US-10-080-943-2 Sequence 2, Appli
25 46 37.4 475 15 US-10-602-268-19 Sequence 19, Appl
26 46 37.4 484 15 US-10-602-268-20 Sequence 20, Appl
27 45.5 37.0 183 17 US-10-425-115-243390 Sequence 243390,
28 45.5 37.0 243 14 US-10-369-493-21737 Sequence 21737, A
29 45 36.6 79 15 US-10-424-599-255885 Sequence 255885,
30 45 36.6 195 15 US-10-424-599-176303 Sequence 176303,
31 45 36.6 307 15 US-10-424-599-171211 Sequence 171211,
32 45 36.6 363 17 US-10-425-115-214263 Sequence 214263,
33 45 36.6 450 17 US-10-425-115-214264 Sequence 214264,
34 44 35.8 51 17 US-10-425-115-307763 Sequence 307763,
35 44 35.8 127 9 US-09-821-883-20 Sequence 20, Appl
36 44 35.8 127 16 US-10-659-295-25 Sequence 25, Appl
37 44 35.8 131 16 US-10-767-701-60435 Sequence 60435, A
38 44 35.8 380 17 US-10-425-115-203649 Sequence 203649,
39 44 35.8 391 15 US-10-282-122A-49957 Sequence 49957, A
40 44 35.8 479 9 US-09-821-883-5 Sequence 5, Appli
41 44 35.8 697 9 US-09-821-883-4 Sequence 4, Appli
42 44 35.8 941 15 US-10-389-566-1845 Sequence 1845, Ap
43 44 35.8 1209 15 US-10-381-882-6 Sequence 6, Appli
44 44 35.8 1683 15 US-10-389-566-778 Sequence 778, App
45 44 35.8 1785 16 US-10-437-963-121407 Sequence 121407,

ALIGNMENTS

RESULT 1

US-10-092-750-54
; Sequence 54, Application US/10092750
; Publication NO. US20030302157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-54

Query Match 100.0%; Score 123; DB 14; Length 22;
Best Local Similarity 100.0%; Pred No. 9, 8e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTRISDMKLADIATWQRNCCPA 22
|||||
Db 1 GTRISDMKLADIATWQRNCCPA 22
|||||

RESULT 2

US-09-864-761-39339
; Sequence 39339, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aemica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24283.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 39339
 LENGTH: 97
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC004126.1
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
 OTHER INFORMATION: SWISSPROT HIT: P46032, EVALUE 4.00e-11
 OTHER INFORMATION: EST_HUMAN HIT: W53019.1, EVALUE 7.00e-27
 US-09-864-761-39339

Query Match 41.5%; Score 51; DB 9; Length 97;
 Best Local Similarity 47.6%; Pred. No. 5.5;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy 1 GTRISDMLKLIADTWORNCPP 21
 Db 72 GSQVSSMLKLAL----QNCPP 88

RESULT 3
 US-10-012-542-141
 Sequence 141, Application US/10012542
 Publication No. US20030044851A1
 GENERAL INFORMATION:
 APPLICANT: Ruben et al.
 TITLE OF INVENTION: 94 Human Secreted Proteins
 FILE REFERENCE: P2029P1
 CURRENT APPLICATION NUMBER: US/10/012,542
 CURRENT FILING DATE: 2001-12-12
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 NUMBER OF SEQ ID NOS: 532
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 141
 LENGTH: 325
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-012-542-141

Query Match 41.5%; Score 51; DB 14; Length 325;
 Best Local Similarity 47.6%; Pred. No. 19;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy 1 GTRISDMLKLIADTWORNCPP 21
 Db 2 GSQVSSMLKLAL----QNCPP 18

RESULT 4
 US-10-115-123-141
 Sequence 141, Application US/10115123
 Publication No. US20030065151A1
 GENERAL INFORMATION:
 APPLICANT: Ruben et al.
 TITLE OF INVENTION: 94 Human Secreted Proteins
 FILE REFERENCE: P2029G30AP1D2
 CURRENT APPLICATION NUMBER: US/10/115,123
 CURRENT FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: PCT/US99/13418
 PRIOR FILING DATE: 1999-06-15
 PRIOR APPLICATION NUMBER: 60/089,507
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089,508
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089,509
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089,510
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/090,112
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090,113
 PRIOR FILING DATE: 1998-06-22
 NUMBER OF SEQ ID NOS: 532
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 141
 LENGTH: 325
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-115-123-141

Query Match 41.5%; Score 51; DB 14; Length 325;
 Best Local Similarity 47.6%; Pred. No. 19;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 GTRISDMLKIADTWORNCPP 21
 ::::|||||
 Db 2 GSOVSSMLKLAL----QNCPP 18
 ::::|||||

RESULT 5
 US-10-257-022-11
 ; Sequence 11, Application US/10257022
 ; Publication No. US20030211499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REDDI, Roopa
 ; APPLICANT: THORNTON, Michael
 ; APPLICANT: BOROWSKY, Mark L.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: KHAN, Farrah A.
 ; APPLICANT: TRIBOUTLEY, Catherine M.
 ; APPLICANT: GANDHI, Ameena R.
 ; APPLICANT: YAO, Monique G.
 ; APPLICANT: SANJANWALA, Madhusudan M.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: NGUYEN, Daniel B.
 ; APPLICANT: POLICKY, Jennifer L.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: SEILHAMER, Jeffrey J.
 ; APPLICANT: WALIA, Narinder K.
 ; APPLICANT: LAL, Preeti
 ; APPLICANT: KEARNEY, Liam
 ; APPLICANT: WALSH, Roderick T.
 ; APPLICANT: LU, Dying Aina M.
 ; APPLICANT: LU, Yan
 ; APPLICANT: GREENE, Barrie D.
 ; APPLICANT: RAUMANN, Brigitte E.
 ; APPLICANT: ARVIZU, Chandra S.
 ; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
 ; FILE REFERENCE: PI-0067 USN
 ; CURRENT APPLICATION NUMBER: US/10/257,022
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: US01/11206
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: 60/195,595
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: 60/196,872
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/199,020
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: 60/200,552
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/202,348
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: 60/203,495
 ; PRIOR FILING DATE: 2000-05-11
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 11
 ; TYPE: PRT
 ; LENGTH: 525
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No: 1563237CD1
 US-10-257-022-11

Query Match 41.5%; Score 51; DB 14; Length 525;
 Best Local Similarity 47.6%; Pred. No. 31;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 GTRISDMLKIADTWORNCPP 21
 ::::|||||
 ::::|||||

Db 258 GSOVSSMLKLAL----QNCPP 274

RESULT 6
 US-09-835-270-2
 ; Sequence 2, Application US/09835270
 ; Patent No. US20020077462A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Rory A.
 ; TITLE OF INVENTION: 33556, A NOVEL HUMAN TRANSPORTER AND
 ; FILE REFERENCE: 381553000800
 ; CURRENT APPLICATION NUMBER: US/09/835,270
 ; CURRENT FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: 60/197,376
 ; PRIOR FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-835-270-2

Query Match 41.5%; Score 51; DB 9; Length 581;
 Best Local Similarity 47.6%; Pred. No. 35;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 GTRISDMLKIADTWORNCPP 21
 ::::|||||
 Db 258 GSOVSSMLKLAL----QNCPP 274

RESULT 7
 US-10-408-765A-904
 ; Sequence 904, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 904
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-904

Query Match 41.5%; Score 51; DB 16; Length 581;
 Best Local Similarity 47.6%; Pred. No. 35;
 Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 GTRISDMLKIADTWORNCPP 21
 ::::|||||
 Db 258 GSOVSSMLKLAL----QNCPP 274

RESULT 8
 US-10-755-889-204
 ; Sequence 204, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company

```
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 204
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-755-889-204

Query Match 41.5%; Score 51; DB 16; Length 581;
Best Local Similarity 47.6%; Pred. No. 35;
Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy 1 GTRISDMLKLIADTWORNCPP 21
Db 258 GSQVSSMLKIAL---QNCPC 274

RESULT 9
US-10-369-493-766
; Sequence 766, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 766
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-766

Query Match 40.7%; Score 50; DB 14; Length 490;
Best Local Similarity 52.6%; Pred. No. 42;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWORN 18
Db 68 TRLSFEQLQKLVSDEWQRN 86

RESULT 10
US-10-425-114-68020
; Sequence 68020, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53113)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68020
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE: Glycine max
; OTHER INFORMATION: Clone ID: 700685326_FLI.pep
US-10-425-114-68020

Query Match 40.2%; Score 49.5; DB 15; Length 332;
Best Local Similarity 43.5%; Pred. No. 33;
Matches 10; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

Qy 1 GTRISDMLKLI-----ADTWORN 18
Db 72 GTSVADLLKLLDSDTNASAWPRN 94

RESULT 11
US-10-424-599-218174
; Sequence 218174, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218174
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39039C.1.pep
US-10-424-599-218174

Query Match 40.2%; Score 49.5; DB 15; Length 338;
Best Local Similarity 43.5%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

Qy 1 GTRISDMLKLI-----ADTWORN 18
Db 78 GTSVADLLKLLDSDTNASAWPRN 100

RESULT 12
US-09-925-301-1295
; Sequence 1295, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1295
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```



```
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1295

Query Match      39.0%; Score 48; DB 9; Length 677;
Best Local Similarity 33.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GTRISDMLKLIADTWQRNCCP 21
Db 552 GLHIDQLVCLVLEAWQKGP 572

RESULT 13
US-10-080-943-4
; Sequence 4, Application US/10080943
; Publication No. US20030073236A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J
; APPLICANT: Tsai, Shih-Chong
; TITLE OF INVENTION: P193 PROTEINS AND NUCLEIC ACIDS, AND USES THEREOF
; FILE REFERENCE: IU99-PCT
; CURRENT APPLICATION NUMBER: US/10/080,943
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/150,266
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-943-4

Query Match      39.0%; Score 48; DB 14; Length 1698;
Best Local Similarity 33.3%; Pred. No. 3e+02;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GTRISDMLKLIADTWQRNCCP 21
Db 1582 GLHIDQLVCLVLEAWQKGP 1602

RESULT 14
US-10-424-599-215390
; Sequence 215390, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215390
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36522C.1.pep
US-10-424-599-215390

Query Match      38.2%; Score 47; DB 15; Length 61;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 4 ISDMLKLIADTWQRNCCP 21
Db 17 ISGGLKNISSPQRIICCP 34
```

RESULT 15

```
US-10-029-386-33933
; Sequence 33933, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33933
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005031.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95
; OTHER INFORMATION: SWISSPROT HIT: Q13075, EVALUATION 1.00e-112
US-10-029-386-33933
```

```
Query Match      38.2%; Score 47; DB 14; Length 203;
Best Local Similarity 53.3%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 7 MLKLIADTWQRNCCP 21
Db 91 LLKKAFLWASGCCP 105
```

```
Search completed: November 11, 2004, 02:43:10
Job time : 26.4953 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-54
Perfect score: 123
Sequence: 1 GTRISDMLKLIADTWQRNCPA 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_79:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	40.7	490	1 JC6038	probable molybdenum
2	50	40.7	490	2 D90727	hypothetical prote
3	50	40.7	491	2 AH0594	probable molybdenu
4	50	40.7	494	2 E85578	hypothetical prote
5	49	35.8	329	2 D89494	2-hydroxyglutaryl-
6	48	39.0	418	2 F96790	hypothetical prote
7	48	39.0	732	2 T43619	secreted protein k
8	48	39.0	732	2 S30060	protein kinase ypk
9	48	39.0	810	2 T93323	hypothetical prote
10	47.5	38.6	3947	2 T52486	ferrichrome sidero
11	47	38.2	498	2 D82093	hypothetical prote
12	47	38.2	517	2 D86264	protein F3f19.6 [i
13	47	38.2	1232	2 A55478	neuronal apoptosis
14	47	38.2	1447	2 T42628	neuronal apoptosis
15	46	37.4	495	2 T27936	hypothetical prote
16	46	37.4	636	2 T10569	probable serine/th
17	46	37.4	2330	1 R81MWV	genome polyprotein
18	44	35.8	127	2 I46269	granulocyte-macrop
19	44	35.8	178	2 S74871	transposase sir124
20	44	35.8	227	2 T04420	ribonuclease (EC 3
21	44	35.8	267	2 T22978	hypothetical prote
22	44	35.8	474	2 B46746	glycine hydroxymet
23	44	35.8	475	2 A33696	glycine hydroxymet
24	44	35.8	1116	2 S77213	DNA-directed DNA p
25	44	35.8	1220	2 A00125	exodeoxyribonuclea
26	44	35.8	2476	2 T34022	zonadhesin - pig
27	43	35.0	98	2 B95323	hypothetical prote
28	43	35.0	334	2 A70101	hypothetical prote
29	43	35.0	377	2 AD2922	altronate oxidored

ALIGNMENTS

RESULT 1

JC6038
probable molybdenum transport protein modF - Escherichia coli (strain K-12)
N;Alternate names: photorepair protein phrA
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: JC6038; H64811; I60601; S32737
R;Grunden, A.M.; Ray, R.W.; Rosentel, J.K.; Healy, F.G.; Shanmugam, K.T.
J. Bacteriol. 178, 735-744, 1996
A;Title: Repression of the Escherichia coli modABCD (molybdate transport) operon by Moc
A;Reference number: JC6037; MUID:96146523; PMID:8550508
A;Accession: JC6038
A;Molecule type: DNA
A;Residues: 1-490 <GRU>
A;Cross-references: UNIPROT:P31060; GB:U27192; NID:G973213; PID:G973219
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64811
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-490 <BLAT>
A;Cross-references: GB:AE000178; GB:U00096; NID:G1786967; PID:AACT3847.1; PID:G178697
A;Experimental source: strain K-12, substrain MG1655
R;Dorrell, N.; Ahmed, A.H.; Moss, S.H.
Photochem. Photobiol. 59, 831-835, 1993
A;Title: Photoreactivation in a phrB mutant of Escherichia coli K-12: evidence for the
A;Reference number: I60601; MUID:94143420; PMID:8310005
A;Accession: I60601
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 151-319 'A', 321, 'R', 323-324 'A', 326-430 <RES>
A;Cross-references: EMBL:X69182; NID:G296234; PIDN:CAA48926.1; PID:G296235
A;Experimental source: K-12 strain C600; cell line JM83/PND01
C;Genetics:
A;Gene: modF; phrA
A;Map position: 17 min
C;Superfamily: Escherichia coli probable molybdenum transport protein modF; ATP-binding
C;Keywords: ATP; DNA repair; GTP binding; nucleotide binding; P-loop; transport protein
F;19-211/Domain: ATP-binding cassette homology <ABC1>
F;36-43/Region: nucleotide-binding motif A (P-loop)
F;276-479/Domain: ATP-binding cassette homology <ABC2>
F;293-300/Region: nucleotide-binding motif A (P-loop)

Query Match 40.7%; Score 50; DB 1; Length 490;
Best Local Similarity 52.6%; Pred. No. 6.8;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;
Qy 2 TRIS--DMLKLIADTWQRN 18
||| : ||| : |||

Db 68 TRLSFEQLQKLVSDWQRN 86

RESULT 2

D90727

hypothetical protein ECs0788 [imported] - Escherichia coli (strain O157:H7, substrain R1)

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: D90727

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

GA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome

A:Reference number: A95629; MUID:21156231; PMID:11298796

A:Accession: D90727

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-490 <HAY>

A:Cross-references: UNIPROT:Q8X941; GB:BA000007; PIDN:BA34211.1; PID:G13360247; GSPDB:C

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs0788

C:Superfamily: Escherichia coli probable molybdenum transport protein modF; ATP-binding

Query Match 40.7%; Score 50; DB 2; Length 490;

Best Local Similarity 52.6%; Pred. No. 6.8;

Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWQRN 18

||| : ||| : |||

Db 68 TRLSFEQLQKLVSDWQRN 86

RESULT 3

AH0594

probable molybdenum transport ATP-binding protein ModF (photorepair protein PhrA) STV081

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AH0594

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moutle, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; Moutle, S.; O'Gaora, P.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AH0594

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05226.1; PID:G16501996; GSPDB:GN00176

C:Genetics:

A:Gene: STV081

C:Superfamily: Escherichia coli probable molybdenum transport protein modF; ATP-binding

Query Match 40.7%; Score 50; DB 2; Length 491;

Best Local Similarity 52.6%; Pred. No. 6.8;

Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWQRN 18

||| : ||| : |||

Db 68 TRLSFEQLQKLVSDWQRN 86

RESULT 4

E85578

hypothetical protein modF [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: E85578

R:Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoculis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85578

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-494 <STO>

A:Cross-references: UNIPROT:Q8X941; GB:AE005174; NID:G12513692; PIDN:AAG55089.1; GSPDB:C

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: modF

C:Superfamily: Escherichia coli probable molybdenum transport protein modF; ATP-binding

Query Match 40.7%; Score 50; DB 2; Length 494;

Best Local Similarity 52.6%; Pred. No. 6.9;

Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWQRN 18

||| : ||| : |||

Db 72 TRLSFEQLQKLVSDWQRN 90

RESULT 5

D69494

2-hydroxyglutaryl-CoA dehydratase, subunit beta (hgdb) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: D69494

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F. Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69494

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-329 <KLE>

A:Cross-references: UNIPROT:O28322; GB:AE000968; GB:AE000782; NID:G2689291; PIDN:AA8892

C:Superfamily: 2-hydroxyglutaryl-CoA dehydratase beta chain

Query Match 39.8%; Score 49; DB 2; Length 329;

Best Local Similarity 52.9%; Pred. No. 6.7;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 TRISDMLKLIADTWQRN 18

||| : ||| : |||

Db 67 TRCCDTLMRLADIWERN 83

RESULT 6

F96790

hypothetical protein F15M4.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 12-Jul-2004

C:Accession: F96790

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F96790

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-418 <STO>

A;Cross-references: UNIPROT:Q9SPW6; GB:AB005173; NID:G6554487; PIDN:AAF16669.1; GSPDB:GN
C;Genetics:
A;Gene: F15W4.19
A;Map position: 1
C;Superfamily: cyclin, A/B/D/E type

Query Match 39.0%; Score 48; DB 2; Length 418;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 TRISDMKLKLIADTWORNC 20
DB 276 TEILDMKLMANTLQNFNC 294

RESULT 7
T43619
secreted protein kinase ypkA - Yersinia pestis plasmid pCD1
N;Alternate names: protein kinase A homolog
C;Species: Yersinia pestis
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T43619; T42850
R;Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Brubaker, J.; Bacteriol. 180, 5192-5202, 1998
A;Title: Structural organization of virulence-associated plasmids of Yersinia pestis.
A;Reference number: Z22578; MUID:98422474; PMID:9748454
A;Accession: T43619
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-732 <HUP>
A;Cross-references: EMBL:AF053946; NID:G2996222; PIDN:AAC62602.1; PID:G2996279
A;Experimental source: strain KIM
R;Perry, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R.
Infect. Immun. 66, 4611-4623, 1998
A;Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia P
A;Reference number: Z22273; MUID:98427122; PMID:9746557
A;Accession: T42850
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-732 <PER>
A;Cross-references: EMBL:AF074612; NID:G3822037; PIDN:AAC69765.1; PID:G3822045
A;Experimental source: strain KIM5
C;Genetics:
A;Gene: ypkA
A;Genome: plasmid pCD1

Query Match 39.0%; Score 48; DB 2; Length 732;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTRISDMKLKLIADTWOR 17
DB 219 GWRCSDTLRLTADSWKQ 235

RESULT 8
S30060
protein kinase ypkA - Yersinia pseudotuberculosis
C;Species: Yersinia pseudotuberculosis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S30060
R;Gal'ov, E.E.; Hakansson, S.; Forsberg, A.; Wolf-Watz, H.
Nature 361, 730-732, 1993
A;Title: A secreted protein kinase of Yersinia pseudotuberculosis is an indispensable vi
A;Reference number: S30060; MUID:93180911; PMID:8441468
A;Accession: S30060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-732 <GAL>
A;Cross-references: UNIPROT:Q05608; EMBL:X69439; NID:G49186; PIDN:CAA49215.1; PID:G49187
C;Keywords: serine/threonine-specific protein kinase

Query Match 39.0%; Score 48; DB 2; Length 732;

Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTRISDMKLKLIADTWOR 17
DB 219 GWRCSDTLRLTADSWKQ 235

RESULT 9
T33323
hypothetical protein C09E8.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33323
R;Du, Z.; Graves, T.; Gibson, A.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of Caenorhabditis elegans cosmid C09E8.
A;Reference number: Z21322
A;Accession: T33323
A;Molecule type: DNA
A;Residues: 1-810 <DUZ>
A;Cross-references: UNIPROT:O9T204; EMBL:AF077529; PIDN:AAC26255.1; GSPDB:GN00020; CESP
A;Experimental source: strain Bristol N2; clone C09E8
C;Genetics:
A;Gene: CESP:C09E8.1
A;Map position: 2
A;Introns: 30/2; 70/1; 118/3; 149/2; 167/3; 293/3; 337/3; 406/3; 449/3; 480/1; 525/3; 5
A;Status: preliminary; translated from GB/EMBL/DBJ

Query Match 39.0%; Score 48; DB 2; Length 810;
Best Local Similarity 35.0%; Pred. No. 22;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTRISDMKLKLIADTWORNC 20
DB 735 GIDVPPLLIIMCNTWIRSCC 754

RESULT 10
T52486
ferrichrome siderophore peptide synthetase - smut fungus (Ustilago maydis)
C;Species: Ustilago maydis (corn smut)
C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 12-Jul-2004
C;Accession: T52486
R;Gentil, G.; Yuan, M.W.; Budde, A.D.; Dohren von, H.; Leong, S.A.
submitted to the EMBL Data Library, July 1996
A;Description: Characterization of the Ustilago maydis sid2 gene encoding a multidomain
synthetic pathway.
A;Reference number: Z26091
A;Accession: T52486
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3947 <GEN>
A;Cross-references: UNIPROT:O43103; EMBL:U62738; NID:G2731632; PID:G2731633; PIDN:AA893
C;Genetics:
A;Gene: sid2
C;Superfamily: non-ribosomal peptide synthetase
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;303-771/Domain: acetate-CoA ligase homology <ACLI>
F;801-870/Domain: acyl carrier protein homology <ACPI>
F;1443-1920/Domain: acetate-CoA ligase homology <ACLI>
F;2540-2991/Domain: acetate-CoA ligase homology <ACLI>
F;3020-3089/Domain: acyl carrier protein homology <ACP2>
F;3054.3611/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 38.6%; Score 47.5; DB 2; Length 3947;
Best Local Similarity 35.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 7; Mismatches 3; Indels 3; Gaps 1;

QY 1 GTRISDMKLKLIADTWORNC 20
DB 536 GEXLS---QVVADEWSKDC 552

RESULT 11

D59093
 Hypothetical protein pXOI-20 - Bacillus anthracis virulence plasmid pXOI
 C:Species: Bacillus anthracis
 C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
 C:Accession: D59093
 R;Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
 J. Bacteriol. 181, 6505-6515, 1999
 A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored
 A:Reference number: A59091; MUID:99445483; PMID:10515943
 A:Accession: D59093
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-498 <OKI>
 A:Cross-references: UNIPROT:Q9X221; GB:AF065404; NID:g4894216; PIDN:AAD32324.1; PID:g489
 A:Experimental source: strain Sterne
 C:Genetics:
 A:Gene: pXOI-20
 A:Superfamily: Bacillus anthracis plasmid pXOI hypothetical protein pXOI-20

Query Match 38.2%; Score 47; DB 2; Length 498;
 Best Local Similarity 72.7%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 LKLIADTWORN 18

| | | | | | | | | |

Db 13 LKLIADTWQKD 23

RESULT 12

D8264
 protein F3F19.6 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: D8264
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.;
 ausen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D8264
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-517 <STO>
 A:Cross-references: UNIPROT:Q9SAD9; GB:AE005172; NID:g4850387; PIDN:AAD31057.1; GSPDB:GN
 C:Genetics:
 A:Gene: F3F19.6
 A:Map position: 1

Query Match 38.2%; Score 47; DB 2; Length 517;
 Best Local Similarity 36.8%; Pred. No. 21;
 Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 RISDMLKLIADTWQNRCCP 21

| | | | | | | | | |

Db 445 RVTEAIKLPEDMKKECCP 463

RESULT 13

A55478
 neuronal apoptosis inhibitory protein - human
 N:Alternate names: NAIP
 C:Species: Homo sapiens (man)
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
 C:Accession: A55478

R;Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird,
 d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneiluk, R.G.; Mackenzie, A.
 Cell 80, 167-178, 1995
 A:Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in ind
 A:Reference number: A55478; MUID:95112344; PMID:7813013
 A:Accession: A55478
 A:Molecule type: mRNA
 A:Residues: 1-1232 <ROY>
 A:Cross-references: GB:U19251

C:Genetics:
 A:Gene: GDB:SWA@: SWA
 A:Cross-references: GDB:120378; OMIM:600354; OMIM:253300
 A:Map position: 5q12.2-5q13
 C:Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane pro
 F;94-110/Domain: transmembrane #status predicted <TM1>
 F;470-477/Region: nucleotide-binding motif A (P-loop)
 F;479-496/Domain: transmembrane #status predicted <TM2>
 F;476/Binding site: ATP (Lys) #status predicted
 F;618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.2%; Score 47; DB 2; Length 1232;
 Best Local Similarity 53.3%; Pred. No. 47;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 7 MLKLIADTWQNRCCP 21

| | | | | | | | | |

Db 479 LKLIADTWQNRCCP 493

RESULT 14

T42628
 neuronal apoptosis inhibitory protein 2 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T42628
 R;Yaraghi, Z.; Diez, E.; Gros, P.; MacKenzie, A.
 Mamm. Genome 10, 761-763, 1999
 A:Title: CDNA cloning and the 5' genomic organization of Naip2, a candidate gene for mur
 A:Reference number: Z22179; MUID:99315342; PMID:10384056
 A:Accession: T42628
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: mRNA
 A:Residues: 1-1447 <YAR>
 A:Cross-references: UNIPROT:Q9QUK4; EMBL:AF102871; NID:g3860228; PID:g3860229; PIDN:AA
 C:Genetics:
 A:Gene: Naip2

Query Match 38.2%; Score 47; DB 2; Length 1447;
 Best Local Similarity 42.9%; Pred. No. 55;
 Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTRISDMLKLIADTWQNRCCP 21

| | | | | | | | | |

Db 517 GSKRTFLKRIAFWASGCCP 537

RESULT 15

T27936
 hypothetical protein ZK622.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T27936
 R;Leimbach, D.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid ZK622.
 A:Reference number: Z20443
 A:Accession: T27936
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-495 <LEI>
 A:Cross-references: UNIPROT:Q23552; EMBL:U39998; PIDN:AAA81102.1; CESP:ZK622.3
 C:Genetics:
 A:Gene: CESP:ZK622.3

A; Introns: 34/2; 97/2; 222/3; 310/3
 C; Superfamily: Caenorhabditis elegans hypothetical protein ZK622.3

Query Match 37.4%; Score 46; DB 2; Length 495;
 Best Local Similarity 43.8%; Pred. No. 29;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 GTRISDMLKLIADTWQ 16
 || :||| :|||
 Db 275 GTSFNLVELIKNTWQ 290

Search completed: November 10, 2004, 14:52:22
 Job time : 6.39623 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-54
Perfect score: 123
Sequence: 1 GTRISDMKLIADTWQRNCCPA 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:
1: uniprot_sprot:
2: uniprot_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	22	Q96JA9	Q96JA9 homo sapien
2	51	41.5	581	Q8IY34	Q8IY34 homo sapien
3	51	41.5	581	Q9P2X9	Q9P2X9 homo sapien
4	50.5	41.1	115	Q9TA93	Q9TA93 sinorhizobi
5	50.5	41.1	444	Q7C0F9	Q7C0F9 anopheles g
6	50	40.7	314	Q6S4V0	Q6S4V0 cotesia glo
7	50	40.7	314	Q6S4V0	Q6S4V0 cotesia glo
8	50	40.7	490	Q829979	Q829979 cotesia glo
9	50	40.7	490	MODF_ECOLI	P31060 escherichia
10	50	40.7	490	Q7AGI1	Q7AGI1 escherichia
11	50	40.7	490	Q7C2Q7	Q7C2Q7 shigella fl
12	50	40.7	490	Q83M03	Q83M03 shigella fl
13	50	40.7	491	Q828A7	Q828A7 salmonella
14	50	40.7	491	Q82Q89	Q82Q89 salmonella
15	50	40.7	494	Q8FJR8	Q8FJR8 escherichia
16	50	40.7	494	Q8X941	Q8X941 escherichia
17	50	40.7	578	Q8BFX9	Q8BFX9 mus musculu
18	50	40.7	578	Q9WU80	Q9WU80 mus musculu
19	50	40.7	936	Q98ST1	Q98ST1 gallus gall
20	49	39.8	329	Q28322	Q28322 archaeglob
21	48	39.0	418	Q9SFW6	Q9SFW6 arabidopsis
22	48	39.0	592	Q7UY6	Q7UY6 rhodospirill
23	48	39.0	732	Q9R112	Q9R112 yersinia pe
24	48	39.0	732	Q9R112	Q9R112 yersinia pe
25	48	39.0	732	Q9R112	Q9R112 yersinia ps
26	48	39.0	832	Q9TZQ4	Q9TZQ4 caenorhabdi
27	48	39.0	832	Q6L4B3	Q6L4B3 solanum dem
28	48	39.0	832	Q6L4B3	Q6L4B3 solanum dem
29	47.5	38.6	1698	Q14999	Q14999 homo sapien
30	47	38.2	344	Q8R3W4	Q8R3W4 mus musculu
31	47	38.2	375	Q6LFM7	Q6LFM7 plasmodium
32	47	38.2	375	CAG25209	CAG25209 plasmodium

32	47	38.2	397	2	Q6D884	Q6D884 erwinia car
33	47	38.2	471	2	Q7FLM6	Q7FLM6 drosophila
34	47	38.2	498	2	Q7CMJ4	Q7CMJ4 bacillus an
35	47	38.2	498	2	Q9X221	Q9X221 bacillus an
36	47	38.2	498	2	AAT28770	AAT28770 bacillus
37	47	38.2	517	2	Q9SAD9	Q9SAD9 arabidopsis
38	47	38.2	678	1	LGP2 HUMAN	Q96C10 homo sapien
39	47	38.2	728	2	Q56921	Q56921 yersinia en
40	47	38.2	729	2	Q85339	Q85339 yersinia en
41	47	38.2	729	2	Q93KQ6	Q93KQ6 yersinia en
42	47	38.2	913	1	CUL7 MOUSE	Q8VE73 mus musculu
43	47	38.2	1160	2	Q8TDZ4	Q8TDZ4 homo sapien
44	47	38.2	1271	2	Q8CH71	Q8CH71 mus musculu
45	47	38.2	1322	2	Q6PB63	Q6PB63 mus musculu

ALIGNMENTS

RESULT 1																
Q96JA9	Q96JA9	PRELIMINARY;	PRT;	22	AA.											
ID	ID															
AC	Q96JA9	Q96JA9	Q96JA9	Q96JA9	Q96JA9											
DT	01-DEC-2001	(Tremblrel. 19, Created)														
DT	01-DEC-2001	(Tremblrel. 19, Last sequence update)														
DT	01-DEC-2001	(Tremblrel. 19, Last annotation update)														
DE	Bcl-XL-binding protein t25 (Fragment).															
OS	Homo sapiens (Human).															
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;															
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.															
OX	NCBI_TaxID=9606;															
RN	[1]															
RP	SEQUENCE FROM N.A.															
RC	TISSUE=Liver;															
EX	MEDLINE=21233069; PubMed=11283018;															
RA	Hammond P.W., Alpin J., Rise C.E., Wright M., Kreider B.L.;															
RT	"In Vitro Selection and Characterization of Bcl-XL-binding Proteins															
RT	from a Mix of Tissue-specific mRNA Display Libraries.";															
RL	J. Biol. Chem. 276:20898-20906 (2001).															
DR	EMBL; AF357524; AAK60628.1; --															
FT	NON TER	1	1													
FT	NON TER	22	22													
SQ	SEQUENCE	22	AA;	2493	MM;	667	FDSC586E5F3B0	CRC64;								

Query Match 100.0%; Score 123; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTRISDMKLIADTWQRNCCPA	22
Db	1	GTRISDMKLIADTWQRNCCPA	22

RESULT 2

Q8IY34	PRELIMINARY;	PRT;	581	AA.	
ID	Q8IY34	Q8IY34	Q8IY34	Q8IY34	Q8IY34
AC	Q8IY34;	Q8IY34	Q8IY34	Q8IY34	Q8IY34
DT	01-MAR-2003	(Tremblrel. 23, Created)			
DT	01-MAR-2003	(Tremblrel. 23, last sequence update)			
DT	01-MAR-2004	(Tremblrel. 26, last annotation update)			
DE	Solute carrier family 15, member 3.				
GN	Name=SLC15A3;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				

JUN 15

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnatone P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.C., Tringwood J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Groumwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.

RESULT 4	
Q9RA93	PRELIMINARY; PRT; 115 AA.
ID	
AC	Q9RA93;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE	ArsR.
DE	ArsR.
OS	Name=arsR; GN
OS	Sinorhizobium sp. As4.
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC	Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX	NCBI_TaxID=102286;
OX	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=As4;
RC	Xu Z.Q., Zhang R.;
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL	-!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
CC	EMBL; AF178758; AAD51845.1; -
CC	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0003700; C:transcription factor activity; IEA.
DR	GO; GO:0006335; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR001845; HTH_ArsR.
DR	InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR	Pfam; PF01022; HTH_5; 1.
DR	PRINTS; PR00778; HTH_ArsR.
DR	SMART; SM00418; HTH_ArsR; 1.
DR	DNA-binding; Transcription regulation.
SK	SEQUENCE 115 AA; 13304 MW; B63B3D6E41EB3F46 CRC64;

```

RESULT 3
Q9P2X9 PRELIMINARY; PRT; 581 AA.
IID Q9P2X9 Q9P2X9
AC Q9P2X9
DDT 01-OCT-2000 (TrEMBLrel. 15, Created)
DDT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DDT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Peptide transporter 3.
GN Name=PTR3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
NCBI_Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUB=Placenta;
RC Ishiiashi K., Imai M.;
RRL Submitted (DEC-1998) to the EMBL/GenBank/DDJF databases.
EMBL; AB020598; BAA93432.1; -
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005215; F:transporter activity; IEA.
GO; GO:0006857; P:oligopeptide transport; IEA.
InterPro; IPRO00109; PTR2.
PFAM; PF00854; PTR2; 1.
SEQUENCE 581 AA; 63559 MW; 56768D2963DBI3D4 CRC64;

```

[illegible]

```

DR EMBL; X69182; CAA48926.1; --
DR PIR; JC6038; JC6038.
DR EcoBASE; EB1628; --
DR EcoGene; EG11677; modF.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; FALSE_NEG.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome; DNA repair; Inner membrane;
KW Molybdenum; Repeat; Transport.
FT DOMAIN 4 235 ABC transporter 1.
FT DOMAIN 261 489 ABC transporter 2.
FT NP_BIND 36 43 ATP (Potential).
FT NP_BIND 293 300 ATP (Potential).
FT CONFLICT 320 320 L -> A (in Ref. 5).
FT CONFLICT 322 322 G -> R (in Ref. 5).
FT CONFLICT 325 325 R -> A (in Ref. 5).
SQ SEQUENCE 490 AA; 54535 MW; 97C220824ED66FB3 CRC64;

Query Match 40.7%; Score 50; DB 1; Length 490;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 TRIS--DMLKLIADTWQRN 18
||| : ||| : |||
DB 68 TRLSFEQLQKLVSDWQRN 86

RESULT 9
Q7AG11 PRELIMINARY; PRT; 490 AA.
AC Q7AG11;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE ATP-binding component of molybdate transport system.
GN OrderedLocusNames=ECS0788;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AP002553; BAB34211.1; --
DR GO; GO:0005524; F:ATP binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding.
SQ SEQUENCE 490 AA; 54479 MW; F7569D0284D1E1C5 CRC64;

Query Match 40.7%; Score 50; DB 2; Length 490;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 TRIS--DMLKLIADTWQRN 18
||| : ||| : |||
DB 68 TRLSFEQLQKLVSDWQRN 86

RESULT 11
Q83M03 PRELIMINARY; PRT; 490 AA.
AC Q83M03;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE ATP-binding component of molybdate transport system.
GN NamesmodF; OrderedLocusNames=SF0544;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin O., Yuan X., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).

```

```

CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE015085; AAN42188.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; Complete proteome.
KW SEQUENCE 490 AA; 54656 MW; ECE2582DF62827AE CRC64;

Query Match 40.7%; Score 50; DB 2; Length 490;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 TRIS--DMLKLIADTWQRN 18
DB 68 TRLSFEQLQKLVSDEWQRN 86

RESULT 12
Q8Z8A7 PRELIMINARY; PRT; 491 AA.
AC Q8Z8A7; Q7C8Q2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative molybdenum transport ATP-binding protein ModF (Photorepair
DE protein PhrA).
GN Name=modF; OrderedLocusNames=STY0811, t2109;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ty2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burdand V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AL627268; CAD05226.1; -.
DR EMBL; AE016841; AAO69726.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR ATP-binding; Complete proteome.
KW SEQUENCE 490 AA; 54656 MW; ECE2582DF62827AE CRC64;

```

```

DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; Complete proteome.
KW SEQUENCE 491 AA; 54745 MW; 3A0FCD9460785F39 CRC64;

Query Match 40.7%; Score 50; DB 2; Length 491;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 TRIS--DMLKLIADTWQRN 18
DB 68 TRLSFEQLQKLVSDEWQRN 86

RESULT 13
Q8ZQR9 PRELIMINARY; PRT; 491 AA.
AC Q8ZQR9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ABC superfamily (Atp bind), molybdenum transporter.
GN Name=modF; OrderedLocusNames=STM0778;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE008732; AAL19716.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR ATP-binding; Complete proteome.
KW SEQUENCE 491 AA; 54696 MW; F25160AA27043BDD CRC64;

Query Match 40.7%; Score 50; DB 2; Length 491;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 2 TRIS--DMLKLIADTWQRN 18
DB 68 TRLSFEQLQKLVSDEWQRN 86

RESULT 14
Q8FJR8 PRELIMINARY; PRT; 494 AA.
AC Q8FJR8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

```

DE Putative molybdenum transport ATP-binding protein modF.
GN NamemodF; OrderedLocusNames=c0836;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:HL / CFT073 / ATCC 700928;
RX MEDLINE=22398234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE016757; AA079309.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000005; HTHArac.
DR Pfam; PF00005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 494 AA; 55044 MW; A90C2E73465DD19F CRC64;

Query Match 40.7%; Score 50; DB 2; Length 494;
Best Local Similarity 52.6%; Pred.No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWQRN 18
|||:|:|:|:|:|:|
Db 72 TRLSFEQLQKLVSEWQRN 90

Search completed: November 10, 2004, 14:50:18
Job time : 31.3057 secs

RESULT 15
Q8X941
ID Q8X941 PRELIMINARY; PRT; 494 AA.
AC Q8X941;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP-binding component of molybdate transport system.
GN Name=modF; OrderedLocusNames=z0930;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / BDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman I.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE005254; AAG55089.1; -.
DR PIR; D90727; D90727.
DR PIR; E85576; E85578.

```

```

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 2.
DR ProDom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 494 AA; 54959 MW; 151FA2B908FA7250 CRC64;

Query Match 40.7%; Score 50; DB 2; Length 494;
Best Local Similarity 52.6%; Pred.No. 27;
Matches 10; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 TRIS--DMLKLIADTWQRN 18
|||:|:|:|:|:|:|
Db 72 TRLSFEQLQKLVSEWQRN 90

Search completed: November 10, 2004, 14:50:18
Job time : 31.3057 secs

```